A Very Corrupt Dialect of the Tongue: Anglicisation and Scotticisms in nineteenth-century Scottish pauper petitions

Groot, Hester

Citation

A Very Corrupt Dialect of the Tongue:
Anglicisation and Scotticisms in nineteenth-century Scottish pauper petitions

Hester Groot

Research MA Linguistics Thesis
Universiteit Leiden
The Netherlands
30 June 2022

Supervisor:
Dr. Moragh Gordon
Second reader:
Dr. Marion Elenbaas

Abstract

The development of late modern Scottish English (1700-1900) is characterized by heavy amounts of prescriptivism. This specifically took the form of a process of anglicisation which pushed the previously high-status Scots language variety into the margins. Scottish linguistic features were proscribed in favour of London-English variants, and numerous studies have attested how the writings of upper- and middle-class Scottish authors took over the higher-prestige anglicised variants at the expense of traditional Scottish forms. How these language ideologies affected the usage of lower-class Scots, though, has long remained underinvestigated, in large part due to an absence of available data. The recent publication of a corpus of lower-class Scottish writing from the nineteenth century—the Corpus of Scottish Pauper Petitions, or ScotPP corpus, which includes pauper petitions written in a variety of Scottish parishes throughout the nineteenth century—offers new possibilities for research into lower-class linguistic developments. Making use of this corpus, the present thesis investigates the occurrence of anglicisation in lower-class written language. I compare the extent to which prescriptivism has affected the ScotPP pauper petitions with the writings of upper- and middle-class Scottish people during this period, drawing on materials the Corpus of Modern Scottish Writing (CMSW). By studying both *overt* and *covert* Scotticisms, drawing respectively on works by contemporaneous prescriptivists and works by modern linguists, an approach is taken that highlights not only the process of language change *from above* through standardisation, but also the *from below* aspect of developments and how they interact and co-occur with the prescriptivism of the era. In doing so, this thesis sheds light on the sociohistorical processes by which anglicisation spread through and affected the language of the nineteenth-century Scottish social classes.

Table of Contents

1. Introduction

The Scottish era of prescriptivism is an important turning point in the linguistic history of Scotland. Much has been written on the period between the sixteenth and the nineteenth century, in which standardisation took firm hold in the Scottish metalinguistic discourse, much as it did in that of their Southern neighbours, the English. Prior to then, there had existed a distinct Scots standard variety spoken in Scotland; however, the changing sociopolitical relations between the Scottish and the English manifested in prescriptivism of a specific kind during this period, namely an anglicisation of language in Scotland. Prominent grammarians and metalinguistic commentators of this time—including famous names such as David Hume and Sylvester Douglas—advocated for speaking and writing the English language in as similar a manner as possible to Standard English, or more specifically, London English. To aid this, a large number of grammars and usage guides were written, in particular throughout the eighteenth century, proscribing so-called 'Scotticisms' and prescribing the Standard English forms that could be used in their stead. And, as has been demonstrated by authors such as Dossena (2002) and Meurman-Solin (1997), Scottish writers of the time heeded the strongly phrased advice of these usage guides, resulting in the Scottish language growing increasingly anglicised.

However, previous research on this topic has been focused primarily on certain specific periods and domains of historical Scottish English. For one, most sociohistorical linguistic research on Scots has focused on periods up to the early eighteenth century, when Scots was believed to have been fully anglicised. However, Dossena (2005) finds that this was primarily the case for written works, and anglicisation was not in fact fully completed in non-print documents. This opens up research areas in the area of non-print documents of the late eighteenth and nineteenth century, whose degree of anglicisation remains largely unexplored. Moreover, while the writings of upper-class Scottish individuals have been studied widely (see e.g. Dossena 2002, Corbett 2013), lower-class Scottish English has remained, by and large, uninvestigated. This is due in large part to the unavailability of suitable data for the investigation of this side of Scotland's linguistic history. Historical linguistic research suffers, as something of an unavoidable rule, from what Labov (1994: 11) coined the *bad data problem*: the availability of preserved materials suitable for linguistic research gets sparser and sparser the further one strays from the present, and possibilities for research are often heavily constrained as a result.

As such, the recent compilation of the Scottish Pauper Petition Letters Corpus (ScotPP) offers a valuable resource and innovative perspective on the linguistic history of Scotland from a heretofore understudied angle. This corpus, compiled at Leiden University in 2022, is composed largely of so-called 'pauper petitions', or letters written by generally lower-class individuals in order to request financial support from their local parishes. This was a core part of the system of poor relief that existed in Scotland at the time, under first the Old Poor Law and, from 1845 on, the New Poor Law; as such, examples of these letters can be found in a large number of archives across Scotland. At the time of writing, the corpus includes letters from both the Highlands and the Perth and Kinross areas of Scotland. This makes the new letters an excellent source for linguistic investigation from a variety of angles, covering such subjects useful to historical sociolinguistics such as regional and temporal variation, but also class-based differences. After all, these letters represent a social class that until now had little representation in the available primary sources for linguistic investigation. Moreover, given the letters' time of writing, in the early- to mid-nineteenth century directly on the heel of the heyday of upper-class prescriptivism, they serve as an excellent resource for exploring how these changes affected the lower classes' language use.

The main question I aim to answer in this thesis is the following: in what ways, and to what extent, was the language of the Scottish lower classes in the early- to mid-nineteenth century affected by the anglicisation of the preceding centuries, particularly as compared to the language of the higher classes? This primary question is investigated along two different routes of analysis. The first of those focuses on the Scottish forms, or Scotticisms, that are seen as *markers* or *stereotypes* of a linguistic community (Labov 1972; or as used in Beal 2004). In the ideology of anglicisation prevalent at the time, these are the forms whose eradication would be taken to herald successful anglicisation of an individual's speech. The second part of this research focuses on typically Scottish forms that were not marked as such, and as such were not the focus of the prescriptivist efforts of linguistic commentators. Their frequency of occurrence, as compared to that of the proscribed former Scotticisms, can provide a valuable insight into the extent to which anglicisation, as an ideologically driven process, focused on just the Scottish forms that were socially marked, as opposed to the full range of Scots linguistic features present in Scottish English at the time.

It is important to foreground that the demarcation of and definition of social class is a challenging subject, especially when imposing class differentiation on historical social structures.

Many factors can play into class status, such as education, income, employment, wealth, and personal background, which are often difficult to quantify together when looking at lesser-known historical individuals. Moreover, the notion of assigning someone a class status runs the risk of imposing anachronistic ideas on societal structure onto history. However, the present research hinges on understanding the social circumstances and structures that influenced language use in the data being studied; the designation of social class allows for a more thorough examination of the connection between sociocultural factors and language change. The process of determining these historical figures' class status for this research is therefore done with care and attention to the fact that in reality, the designation of social status was not as categorical as with the methods used in this study. The specifics of this process will be elaborated on in Chapter 3.

Aside from answering the above questions, it is a noted goal of this work to draw attention to the possibilities afforded us by the new data represented in the ScotPP corpus. A key part of not just historical linguistics, but also the field of linguistics as a whole, is making workable data available for further research among the scholarly community, in order to further our understanding of new aspects of language to the extent that the data allows. Through this thesis, I therefore aim to show just one of many possibilities for using this newly retrieved source of historical linguistic information, in the hope that more roads of scholarly research will be travelled in the investigation of these pauper petition letters, given their value as a source of language history 'from below' (a concept which Chapter 2 will expand on).

This research project has been organised in the following way: Chapter 2 consists of a literature review, which first describes the field of historical sociolinguistics, and research that has been done on the areas of prescriptivism and standardisation that form the backdrop of this study. Then, it will delve into the specifics of the Scottish historical and linguistic situation of the early- to mid-nineteenth century, and present the primary source materials that will serve as the grounds for this study. These materials consisted of both lower- and higher-class correspondence from this period, as well as data from a variety of usage guides published in the preceding decades. Chapter 3 expands on the methodology of this study, and Chapter 4 outlines the results found in the investigation of the correspondence data. Chapter 5 provides an analysis of the results for the various sets of letters, taking into account extralinguistic factors such as social class, region, and discursive function of the correspondence; it also considers the non-proscribed forms that occur throughout the data, and leads into a discussion of how the results of the study, both in what *is*

present in the data and what *isn't*, might be able to reflect the course of modern-day Scots and Scottish English. Chapter 6 presents the conclusions that can be drawn from its findings and possibilities for future research.

2. Theoretical background

2.1 The field of historical sociolinguistics

The discipline of historical sociolinguistics originates in research from the 1980s and 90s, spearheaded by scholars such as Suzanne Romaine, James Milroy, and the collaborative work of Nevalainen and Raumolin-Brunberg. In the decades following, the field has established itself as an area of linguistics that involves, among other things, "[r]econstructing how language changes diffuse socially" (Nevalainen & Raumolin-Brunberg 2003: 10) and establishing "a broad picture of the social context in which the language varieties under investigation were used" (Auer et al. 2015: 5). The discipline is built on the assumption of the *uniformitarian principle*, which states, in the words of Labov (1972: 275), that "the forces operating to produce linguistic change today are of the same kind and order to magnitude as those which operated five or ten thousand years ago". As such, historical sociolinguistics can draw on modern methods and principles from sociolinguistics to explain and trace historical linguistic developments.

Working off the basis of the aforementioned uniformitarian principle, and with the documents that they have at their disposal, historical sociolinguists work to investigate language as it existed in a given historical period. The focus here lies on synchronic as well as diachronic perspectives on a language's history, and on the social factors that often underlie a language's development. These factors can include aspects like social networks and interaction patterns, but also forces like language attitudes and standardisation. The latter two are of particular interest to the present study.

Historical sociolinguistics primarily makes use, often by necessity, of written data, as spoken material is not available for the vast majority of the history of language. In modern sociolinguistics, oral language is generally the preferred source material, since it is (in many cases) more spontaneous and less governed by prescriptivist language rules than written language is (though this is a generalisation that, of course, knows many exceptions). Historical sociolinguistics has therefore long sought, much like the broader discipline of historical linguistics, the written sources that most closely resemble the more spontaneous side of spoken language. Today, that approach has become somewhat more multifarious, with another goal being to find as wide a range of data types as possible—i.e., to have data that are representative of a given society in many different aspects along a cline of informal (spontaneous), but also, formal (more closely monitored) language.

Written sources, after all, are not all cut from the same cloth. Indeed, a great deal of variation exists between various text types, based on register, communicative intent, and writer characteristics. Koch and Oesterreicher (1985) make the distinction between the *medium* of a text—a strictly binary categorisation of phonic versus graphic—and the *conception*, which categories texts based on their communicative strategies. This conception is characterised as a continuum, and the authors term the poles of this continuum the 'language of immediacy' (or *Sprache der Nähe*) and the 'language of distance' (or *Sprache der Distanz*). These represent, respectively, the prototypically oral texts (characterised as informal, dialogic, unplanned) and the prototypically written (monologic, formal, planned). Not only written texts can fall anywhere on this conceptual continuum, as the same goes for spoken texts, such as political speeches or sermons.

Within this conceptual framework, the discipline of historical sociolinguistics often seeks written sources that fall on the side of the language of immediacy, being more unplanned, and therefore closer to the natural spontaneous speech that is in many cases the object of study. Such sources have been argued by many historical sociolinguists (e.g. Nevalainen & Raumolin-Brunberg 2003, Elspaß et al. 2007) to be the best way of accessing authentic historical language, as close to the way it was spoken as possible. This brings researchers to text types known as 'ego-documents', a term for texts written in the private sphere "in which an author writes about his or her own acts, thoughts and feelings" (Dekker 2002: 14)—a term that covers the genres of interpersonal letters, travelogues, diaries, and similar text types (though of course, conceptual variation occurs within these text types, too).

Ego-documents come with several important advantages compared to other written materials. Firstly, these documents were generally intended to be read by a considerably smaller audience than traditional written works such as literature, legal documents, etc.; they were often crafted either in personal correspondence with a single individual, or, as in the case of personal diaries, truly intended for the writer's eyes only. This lends the language a certain unmediated quality, less carefully constrained and editorialised than literary works would be. In this view, ego-documents are closer to Koch and Oesterreicher (1985)'s idea of the language of immediacy. In terms of the end goal of investigating language as close to the spoken domain as possible, ego-documents are in many historical cases the best available route for researchers.

The use of ego-documents also has another benefit when it comes to trends in the field of historical linguistics. One problem that traditionally characterised many language histories was the tendency to focus on processes of language change and standardisation as teleological; Watts (2012: 585) wrote that language historiography has traditionally been grounded on "an implied teleology [...] that standard languages are the only valid objects of study for a language history". This practice leads to the erasure of non-standard language varieties in favour of the standardised forms we later came to know, and makes for incomplete language histories. After all, it leaves a considerable portion of a language's historical variation and speaker base unrepresented.

As a result, a focus that has been followed increasingly by historical sociolinguists in past years (e.g. Elspaβ et al. 2007) is that of the *language history from below*. Elspaβ et al. (2007: 5) define the term 'from below' as "the social ranks below the highest social class and to texts representing everyday language that could thus be considered as below formal registers such as the language of literature". This change of focus within historical sociolinguistics brings the writing of the previously understudied lower classes to the forefront, and with it, the ego-documents that best reflect their forms of language. What makes ego-documents special in this regard is that they were generally produced by a far broader range of people in society than literary works, for example, the writing of which was usually limited to the elite. Research into standardisation and prescriptivism processes no longer primarily has to focus on the high prestige, upper-class materials that were the object of much of previous work into this area.

A key part of studying the language of lower classes is understanding the processes that allow some of their linguistic patterns to persist and grow, while—deliberately or non-deliberately—consigning others to memory. This relates closely to the role that language attitudes and ideologies play in the development of a language. Such factors often involve social attitudes regarding class and class-tied linguistic markers, as well as markers of other social characteristics such as geography, gender, etc.—characteristics that may come with their own ideological perceptions, which are large drivers in the processes of standardisation and language change. In the following section, I will discuss how the field historical sociolinguistics has focused on standardisation and prescriptivism in the past, how the *language from below* approach can be applied to this aspect of the field, too, and how these subjects are important to the study of the language of the historical Scottish lower classes.

2.2 Perspectives on standardisation and prescriptivism

Standardisation has been the subject of much historical sociolinguistic research across languages. It has historically been a far-reaching, primary factor in determining the course of language change and variation, and the development of many of today's most spoken languages is affected by standardisation processes, which can be orchestrated with varying levels of deliberateness. Milroy (1994) describes *standardisation* as an umbrella term, which is used to describe complex processes of making choices between various sets of linguistic variants. At its basis, though, the process involves the heightening of the status of one language variety over that of another; this process develops in the direction of homogenisation of language use in a particular area.

Millar (2012) summarises this process as follows: "[o]ne variety of a language becomes associated with power and prestige; eventually use of other varieties, particularly in writing, becomes highly marked" (44). This development can take different forms; today, he writes, this status-raising of one variety to the deferment of another is generally a planned process, referred to by Joseph (1987) as *engineered standardisation.* In this scenario, an individual or group of individuals deliberately raises the status of a language, including its orthography and grammatical and lexical structure. In the past, by contrast, standardisation generally developed through a process of *circumstantial standardisation*, in which there is no conscious and/or deliberate plan to raise the status of the variety that becomes the standard. This scenario involves a semi-conscious and, at the very least, improvised push towards the raising of a variety's status.

Language varieties developing along the lines of these respective forms of standardisation are associated with different linguistic processes. Engineered standardisation is often enacted through a process of codification—that is, the establishment of certain language norms as rule, typically in the form of officially-sanctioned dictionaries and grammar books. Moreover, codification often occurs as an element of governmental language policy and planning efforts to direct the development of linguistic standards. Circumstantial standardisation, on the other hand, often occurs due to the lowering of a language variety's prestige value as a result of broader societal developments. Millar (2012: 66) describes the development of Scottish English near the end of the seventeenth century in terms of circumstantial standardisation and the *dialectalisation* of Scots—i.e., the variety being subsumed under another language, in this case Standard English. As we will see, elements of both engineered and circumstantial standardisation can be found in the history of Scottish English throughout its history.

The role of standardisation in the history of Scottish English has been well-attested in linguistic research. Particularly the process of anglicisation that occurred within its standardisation has received a great deal of scholarly attention; its course, which reared its head in the early sixteenth century and carried over through the following centuries, was discussed already in the early twentieth centuries by pioneers such as Bald (1927). Before delving into the current state of scholarship on the topic, a brief sketch of Scotland's sociohistorical and (meta)linguistic background will help to understand the linguistic developments that are connected to the anglicisation process.

2.3 Scotland, Scottish English, and anglicisation

The linguistic landscape of Scotland has a complex history. Today, the main language varieties that are in use include Scots, Scottish Gaelic, Scottish Standard English, and Standard English. Of these languages, Scottish Gaelic is an outlier in the sense that it does not stem from Old English, having developed from Old Irish instead. As for Scots, Scottish Standard English, and Standard English, these are varieties that each developed from Old English. The overarching term for these varieties of English spoken in Scotland is, fittingly, Scottish English, though the term 'variety' itself is a loaded one, given the debates on whether Scots itself can be termed a language in its own right. This thesis will use the term 'language variety' throughout, where relevant. Moreover, while the present focus is on the interaction between Scots, Scottish Standard English, and Standard English—which have throughout history often formed more of a continuum than a series of entirely discrete languages—the context of Scottish multilingualism is relevant to keep in mind when discussing its historical linguistic landscape.

Throughout their eventful histories, Scots and English have frequently converged and consequently diverged again over the centuries. Devitt (1989) finds that the sixteenth century was a turning point for Scots-English; the language variety had been on an ongoing rise in status and usage before then, which peaked during this century, but the period also marks the beginning of the variety's fall in public opinion. Where Scots-English had long been the standard variety in Scotland, that position was gradually taken up by Standard English, instead. This development is closely connected to the sociopolitical developments between Scotland and England of this period. The Union of the Crowns in 1603 united Scotland and England into a single kingdom—though rivalry between the two continued—and heralded the onset of a process of anglicisation of Scots-

English that was closely tied to this event and the shifted socio-political circumstances (e.g. Dossena 2002: 105). Those shifting socio-political developments continued through the seventeenth century, with a turning point being the Union of Parliaments in 1707. This event, wherein the government of Scotland was moved to London, had a strong impact on the sense of Scottish status that was imagined both in Scotland and outside of it; Dossena (2002: 106) writes that Scotland gained a "sense of marginality", both socio-politically and sociolinguistically.

This process is illustrated well by the following quote, which John Sinclair (1782), soon to be further discussed later in this chapter, wrote in the introduction to his influential usage guide on the linguistic situation of Scotland:

> "It ought alſo to be obſerved, that it is very natural for an inferior kingdom to imitate the manners and language of a wealthier and more powerful neighbour: a circumſtance ſtill more to be expected, when both nations came to be governed by the ſame King, who ſeldom viſited Scotland, and who would not offend the prejudices of his new ſubjects, by permitting any other language to be made uſe of at his court, than that of England."

The process of anglicisation of the Scots language, in which many of its distinctly Scottish features were side-lined in favour of Standard English equivalents, did not start only after these events at the turn of the eighteenth century. Marjorie Bald wrote a series of works in the 1920s in which she found that Scottish anglicisation had its origins in the sixteenth and seventeenth century, perhaps with the invention of printing in Scotland, which led to the spread of English printed texts and with them English orthography throughout Scotland. As such, by the time the Union of Parliaments rolled around, the process of anglicisation had already been set in motion; however, the eighteenth century took this process up with a renewed vigour. Given the aforementioned sense of marginality that the preceding political events had lent the people of Scotland, many people, generally those from the upper echelon of society, saw social benefits in taking up the speech patterns of their Southern neighbours, and generally shedding all outward signs of Scottishness in order to partake in the preferred ranks of English society. The sentiment towards Scots of the time is reflected well in the words Alexander Geddes (himself, despite the below quote, a supporter of using Scots more broadly):

"Few perſons of genius and learning will be inclined to write in the Scottiſh dialect; and if any were inclined, they could not look for encouragement or imitators. Men will ever follow thoſe purſuits that lead to riches or fame; and Scottiſh compoſition, either in proſe or poetry, will neither fill the writer's purſe nor emblazon his reputation." (Geddes 1792: 404)

Numerous studies have been conducted on the language of the upper-class Scots of this time. Cruickshank (2011) investigated the language, and anglicisation thereof, of James Duff, Second Earl Fife, in the late eighteenth century; Corbett (2013) discusses the orthographic practices of prominent eighteenth-century Scottish writers Allan Ramsay and Robert Burns. There is also Dossena (2002)'s work on the subject of code-switching between Standard English and Scots, which covers the correspondence of upper-class Scottish authors such as the late nineteenth-century author Robert Louis Stevenson, as well as comparing the linguistic picture of Stevenson's writing to fictional accounts of Scots from the same century.

An important source on the anglicisation of written Scots is Meurman-Solin (1997), who uses the writings and correspondence of primarily upper-class Scottish people to trace the course of said anglicisation. Drawing on materials from *The Helsinki Corpus of Older Scots* and the writings of individual Scots such as Thomas Stewart (1668-70), Andrew Hay (1659-60), and William Cunningham (1673-80), Meurman-Solin is able to trace the course of anglicisation from its initial beginnings, which in writing took hold in the early decades of the seventeenth century (1997: 16). In those early decades, the linguistic developments in Scotland are complex to trace, as the beginning stages of anglicisation overlap and interact with lingering traces of the process of de-anglicisation that Scots underwent in the centuries prior. During those centuries, the higher prestige value and sociocultural status of Scotland had led to the divergence of the language from Southern English; correspondence ego-documents offer Meurman-Solin a source for tracing the development of the ensuing convergence.

A number of findings stand out in Meurman-Solin's picture of that early Scots-English convergence. In the Scots language variety, or its written form at least, it appears to have been men who were the first adopters of English spelling variants such as *-ed*, *wh-*, and *they* (over the Scottish variants *-it/-yt, quh-*, and */thai(i)/thay* (Meurman-Solin 1997: 18). Moreover, Meurman-Solin

(1993) finds that an important factor that determines the degree of anglicisation or lack thereof in a letter is the relationship between writer and addressee. She notes that Scottish features are used primarily in letters written to social inferiors; that relationship is labelled 'intimate down' in her materials, and includes, for instance, letters written by parents to children, or husbands to spouses. By contrast, letters labelled 'intimate up', and non-intimate letters written to socially superior recipients featured markedly fewer Scottish features.

Such factors have been highlighted by Biber (1994), who proposed a framework for determining register and linguistic choices as determined by social and situational circumstances. Biber divides a variety of possible factors influencing a linguistic situation into *social factors* and *situational factors*, including under the first category factors such as gender, age, social position and social networks. Under the situational factors, Biber classifies factors including relations between addresser and addressee, communicative purpose of the document, and channel and topic of the text. Each of these factors might be of influence on the linguistic choices made by a speaker or writer, which in the case of Scots and anglicisation results in more retained Scottish features when the distance and hierarchy difference between participants was smaller.

A primary motor of the anglicisation of the Scots language was the phenomenon of the usage guide. The eighteenth century saw a stunning boom in the number of usage guides being written, following the popularisation of the genre among the English neighbours with Robert Lowth's *A Short Introduction to English Grammar* (1762). The usage guide was, and remains to this day, different from the grammar because its focus lies not exclusively on grammar, as is the case with its eponymous counterpart; instead, usage guides often focus on a variety of aspects of language, including grammar but also touching on topics like lexicon, register, etc. Today, the genre remains popular, and new usage guides continue to be published; back in the eighteenth-century heyday, they epitomised and strengthened an age of prescriptivism that would cause previously unseen levels of anglicisation of the Scots language.

Eighteenth-century usage guides were primarily aimed at the upper classes of Scottish society, and scholars such as Dossena (2005) have highlighted their marked effect on the anglicisation of those target audiences' language use. It was not only the upper classes of Scottish society who had use for the popular usage guide genre, though. Literacy was a widespread quality in late modern Scotland, the extent of which has been much discussed and often exaggerated by historians and historical linguists. The middle classes had a need for literacy in order to be able to

read the laws that applied to their businesses, trade, and commerce; but even lower classes were widely exposed to some degree of literacy. The reformed kirk established parish schools throughout Scotland, which taught reading and writing with one of the primary purposes being to allow people to acquaint themselves with the vernacular Bible (an English translation, not a Scots one).

While this policy certainly allowed the Scottish lower classes to acquire a degree of literacy that exceeded that of many of their foreign contemporaries, the extent of that literacy has been questioned and deconstructed by Houston (1985). Houston finds that the quality of these parish schools varied, and the foreign education that poorer children in the seventeenth and eighteenth centuries were exposed to was generally brief and incomplete.

It is then an interesting question to what extent those lower classes' language use was actually affected by the usage guides. Many of those were not aimed at them in the first place; Hume (1752)'s list of Scotticisms, for instance, was described by Cruickshank (2011: 19) as "more of an aide memoire for himself and for those adult Scots already established in Edinburgh society". But as similar works continued to be published with increasing frequency, authors began to appear who did explicitly target non-upper class audiences. Mitchell (1799), for instance, aimed his book of Scotticisms at classrooms of urban middle-class students, and thereby a wider, less socially prestigious audience than his predecessors (Cruickshank 2011: 30). As such, there existed some basis, at least in prescriptivists' intentions, to get through to the lower classes of Scottish society and anglicise the patterns of speech of these individuals, too.

While a number of works, as discussed above, have investigated the degree to which the language of the Scottish upper classes underwent the process of anglicisation, it is difficult to determine how strong the effect of this prescriptivism and the increasing popularity of usage guides was on the language of the lower classes. Besides the topic being a challenging one to measure concretely—as will be further expanded upon later—there is the simple fact of availability of data. The Labov-termed *bad data problem* plagues the field of historical (socio)linguistics, and particularly lower-class data often presents a challenge to get hold of: the materials that are suitable for such research—ego-documents, preferably private correspondence, or perhaps personal journals or travelogues—generally skew overwhelmingly towards those produced by higher-class individuals. As such, while the effect of efforts towards anglicisation on the language use of the Scottish upper class have been studied at some length by historical sociolinguistics, the question

of how the usage guide influx affected the language of the lower classes has thus far remained uninvestigated, primarily because the data that would make such a study feasible was simply unavailable.

Therefore, the Corpus of Scottish Pauper Petitions (also known as the ScotPP corpus [2022]), built at Leiden University by Dr. Moragh Gordon, Dr. Jelena Prokic, Hester Groot, and Alma Strakova, offers a unique and new insight into the language of the Scottish lower classes of the nineteenth century. Below, I will discuss firstly the nature of pauper petition letters and why they form such a promising source for investigating this question, and secondly the background, contents, and setup of the corpus.

2.4 Pauper petition letters as a source for lower-class Scottish English

The materials that this thesis will introduce into the picture of Scottish English as spoken in the mid-nineteenth century are pauper petition letters, which stem from the system of the Poor Law, specifically the New Poor Law that was implemented in Scotland in 1845. Pre-1845, Paterson (1976: 171) writes that each Scottish parish held responsibility for its own poor. People's right to belong to a parish was determined by settlement, i.e., having been born there or having resided and worked there for three years. A claimant had to be 'destitute' and 'disabled' in order to qualify for poor relief; if these conditions were met, the individual would be placed on the parish poor roll and granted financial aid as a pauper. This system was changed in 1845, with the implementation of a new Act, and the system of poor law that followed it became known as the New Poor Law. This Act brought about a number of changes, including the requirement that parishes provide levels of support sufficient to the needs of their claimants; the hiring of paid inspectors to oversee the poor relief management in parishes; and the establishment of a Board of Supervision (Jones & King 2016).

An important component of the system of poor relief, not just in Scotland but in the whole of Britain at the time, was that claimants would write letters to their parishes in order to outline their situation and establish themselves as 'destitute' and 'disabled', in the hopes of being admitted to the aforementioned poor rolls. These letters are known as *pauper petition letters*, or PPLs, and they have in recent years proved a valuable source of (often) lower-class writing from a period of which not much of such writing has survived. Previous work has been done on PPLs in England,

with the Language of Artisans and the Labouring Poor, or LALP, project being carried out over the past decade (cf. Auer and Fairman 2013).

Until now, little research, particularly linguistic research, has been done on Scottish PPLs. Historians Jones and King (2016) have used a selection of Tongue PPLs to investigate the changing relationship between pauper and parish in the mid-nineteenth century; Barclay (2017) studied the relationship between masculinity and begging in Scottish PPLs. But on the whole, this remains an area with a lot of interesting unexplored research directions, including the historical linguistic questions of language variation and change.

What makes PPLs particularly interesting for historical linguistic research is that they potentially represent the writing of a social class whose work traditionally has been preserved the least: the writing of the lower classes. History has a tendency to preserve writing by upper classes, who of course generally also had the most access to education and thus the levels of literacy it would require to create said writing in the first place. As a result of these factors, the writing— and, by extension, language use—of lower-class populations throughout history has been far less documented than that of higher classes.

The discovery of pauper petition letters thus offers a rich source for historical linguistic research into lower-class language use, thereby filling the research gap described above. This should be qualified a little, as PPLs were often not in fact written by the petitioners themselves, who instead enlisted help from scribes. These scribes, who might be known to the petitioners as officials, family members or acquaintances, generally had a level of education and literacy that exceeded that of the petitioner, which influenced their writing practices and knowledge of standard grammar rules. The resulting writing is, consequently, removed a step from the unmediated spoken forms used by the petitioner and/or scribe in a way that should be accounted for in any analysis of this data.

This phenomenon of differing authorship is clearly visible in the PPLs preserved from the parish of Perth, a town in the central lowlands of Scotland. As with a number of Scottish parishes, the letters written by petitioners to the Perth parish have been preserved, and form a primary source for a snapshot of Scottish English around the mid-nineteenth century. A number of the PPLs written in Perth were written by the same scribe, a man named Alexander Mackenzie. These letters are often signed by Mackenzie himself, making them easily identifiable. Other times, the name signed at the bottom is that of the petitioner themself, but includes a so-called X signature, in which

the petitioner leaves an X in place of a written signature at the bottom of a document written up by someone else. These X signatures can therefore also be an indication that a letter was not written by the petitioner whose words it ostensibly portrays.

However, many letters were also penned by the petitioners addressing their own request to the parish. While it is not always possible to say this for certain, indications that this is the case can include the use of first person pronouns—often crossed out in favour of third person forms, which were the norm in PPLs—which form a clear indication that the writing was done by the person whose name is on the page. These letters represent a societal group that has thus far been little-documented in historical linguistic research.

Structurally, the ScotPP corpus consists of two sets of sources. One set consists of letters, reports, and additional documents from Tongue, a settlement on the far north coast of the Scottish Highlands, all dating from around the early 1850s. Around the mid-nineteenth century, the majority of Tongue's population consisted of crofters, a Scots term for smallholders who made a living off parcels of land let to them by estate managers. Following population growth, the land was further subdivided in the eighteenth century, leaving the parcels entirely inadequate for sustaining their crofters, with food scarcity and famine always in close reach (Jones & King 2016: 80). Minister of the parish Hugh McKay MacKenzie wrote in 1841 that "The general standard [of living] is … wretchedly low. No doubt a few of [the residents] are comfortable, but the generality seldom can rise above the commonest necessaries of life" (*Statistical Account,* xv: 177; as quoted in Jones & King 2016: 82). Those residents who could not maintain their own upkeep are the ones represented in the dataset, who wrote to the parish out of necessity to petition for poverty relief.

The other set of documents stems from Perth, a larger parish in the central lowlands of Scotland. These letters were written around the early 1820s. Perth was, in this period, an industrial town; *Perth Academy* had been founded there in 1761, and brought with it a major industry. Education levels were high throughout the nineteenth century, according to accounts by e.g. Anderson (1983). The literacy rates in the Scottish lowlands have historically been described as notably high for the time, with male literacy lying around 65 percent in the mid-nineteenth century, and female literacy at 25-30 percent (Houston 1985: 56-62). Anderson (2018: 100) notes that the Industrial Revolution, which hit Perth heavily, likely worsened overall literacy levels, not just in the lowlands but in Scotland in its entirety. However, he concludes that on the whole, by the early 1800s, literacy had generally permeated all levels of the lowlands' social structures, and "the

ability to read was broad enough to support the beginnings of a tradition of working class self-education and self-improvement" (2018: 100). It is then perhaps surprising that the proportion of pauper petitions in Perth that were written by scribes was markedly higher than that of Tongue, where the majority appears to have been written by petitioners themselves.

2.5 Comparative material: the CMSW letters

In order to be able to contextualise the data of the Perth and Tongue PPLs, it is necessary to include comparative material for contextualisation purposes. If we want to understand how the language of the Scottish lower classes developed and differed from the prevailing, top-down ideas of historical Scottish English that historical linguists have drawn up in past decades, those prevailing linguistic pictures need to be held side by side with the new PPL data. As such, a second set of data is necessary, one which includes written materials from the same period of Scottish history— that is, covering a period from approximately the 1820s to the 1850s—from the higher classes, both middle and upper, of Scottish society.

Luckily, a number of different corpora have been built in past years that feature correspondence from this demographic. These include Marina Dossena's Corpus of Nineteenth-century Scottish Correspondence (19CSC); Anneli Meurman-Solin's Corpus of Scottish Correspondence (CSC); and John Corbett et al.'s Corpus of Modern Scottish Writing (CMSW). Each of these corpora features an amount of personal correspondence written between identifiable individuals during the nineteenth century, including primarily higher-class individuals (whose social class is often discernible by their profession).

The corpus selected to serve as a supplier of a comparative dataset was the CMSW, built at the University of Glasgow in 2007. This corpus, which covers the period of Modern Scots from 1700 until the present day, was (as is stated in its description) specifically built with the goal of helping research trace the course of anglicisation of the Scottish language throughout this period. Part of the corpus, in fact, consists of material written by orthoepists, or language commentators, highlighting the trajectory and talking points of Scottish prescriptivism throughout this period. Primarily, though, the corpus includes a variety of personal correspondence, among which around 50 letters dated between 1820 and 1860. The authors of these letters include people with professions such as publisher, author, physician, and, in the case of the included David Livingstone, missionary and explorer. These professions were taken as general markers of social

class and status, as highly skilled roles that both required a significant amount of scholarly education and, most likely, paid a good deal, too.

As ego-documents, the letters included in the CMSW are more varied than those in the ScotPP corpus. Where the latter consists near-entirely of pauper petition letters, a genre of correspondence governed by rigid formulaic rules and a clearly defined hierarchical writer-addressee relationship, the letters included in the CMSW show represent a considerably larger range of purposes, registers, and social relations. As Dossena (2013: 103) notes, materials can include familiar letters (such as those written by a Mary Mather to her sister Elizabeth Buchanan), letters between friends, but also letters dealing with professional matters such as job applications. Moreover, the document length in the CMSW corpus varies considerably; although the greatest outliers are not included in the present dataset, as they do not relate to correspondence, the shortest letter included in this study consists of just 42 words, while the average letter length lies around 500 words.

The difference in writer-addressee relations in the CMSW as compared to the ScotPP corpus, as mentioned above, is interesting in terms of class aspect and linguistic content. While in the ScotPP corpus, the writers are generally of a lower class, writing up to individuals higher in social rank and status, that is not necessarily the case for the CMSW letters. In the latter, writer and addressee generally occupy a similar social status, though letters can still contain requests, much as they do in the ScotPP. This difference in class dynamics between the corpora, though, is of interest to the linguistic strategies used by letter writers. Social distance, accommodation, and hierarchy may all play a role in determining the linguistic forms that are used throughout the texts. Here, again, Biber (1994)'s framework of social and situational factors influencing register is of relevance again; particularly the parameters he outlines pertaining to *Relations between Addressor and Addressee* and *Purposes, Intents and Goals* are areas wherein the CMSW and ScotPP letters diverge, and will be of interest in the discussion of this study's results.

2.6 Usage guides as a source for marked Scotticisms

Having detailed the materials that will be used for comparison in this study, the next step is to establish a framework for how to investigate the feature of interest in these letters—that is, the anglicisation of the respective corpus datasets. Determining how 'Scottish' a text is can be a dicey subject; Corbett (2013: 83-4) warns against the notion of a text's 'Scottishness' as an empirical

diagnostic or index. Often, a text may be deemed 'Scottish' based on the occurrence of a single Scots lexeme or orthographic sign, when that sign may be a unicum in an otherwise Standard English text. Similarly, the presence of a singular Standard English form cannot be taken as a sign of anglicisation at face value; it is important to weigh context and co-occurrence in an assessment of said anglicisation.

However, the occurrence of Scottish lexemes or orthographic signs can shed light on how a text falls within a societal, metalinguistic perspective on 'Scottishness'. To this end, a valuable source for deciding on variables and markers of socially marked Scotticisms can be the wealth of usage guides published around the late eighteenth and early nineteenth century, whose contents strongly reflect the prevailing societal ideas surrounding Scottishness and the stereotypes associated with them. What is salient here is that this relates specifically to features that could be socially and culturally recognised as Scotticisms—in short, what Labov (2001) described as the stereotypes, as opposed to the markers, of a linguistic 'style'. This distinction relates to what Aitken (1979) terms *covert* versus *overt* Scotticisms, and it is a distinction that is very relevant here. Overt Scotticisms are those that are marked in the public discourse as being Scottish, and may therefore for instance occur in literary representations of Scottish people in which their Scottishness is being emphasised; these features are the ones that will often be remarked on by authors of usage guides. Covert Scotticisms, on the other hand, are those that are in use in the day to day but need not be markedly Scottish in the perception of speaker and listener. (More on this will be said in Chapter 3). Covert Scotticisms are those that are used unselfconsciously.

In choosing usage guides to investigate the occurrence of marked Scotticisms as compared to the use of Standard English forms, the Scotticisms that are highlighted are those that were, at the time, considered highly marked and overt. The forms commented on and proscribed in usage guides are those that were, by definition, recognisable to commentators as Scottish forms, and in need of eradicating. As such, making use of usage guides for the analysis of anglicisation in nineteenth-century Scottish English gives a perspective that focuses, by necessity, on the Scottish forms that were socially marked at the time. This makes for a valuable perspective because, of course, the anglicisation of the Scots language was also a very socially and ideologically driven process; the metalinguistic angle is essentially built into the process.

This approach can be complemented with a further investigation of the presence of *covert* Scotticisms in the language of the various social classes of nineteenth-century Scotland. This

additional perspective uses modern descriptive writings on Scotticisms (a term that no longer carries the derogatory tone it did in the prescriptivist age of Scotland) to investigate the presence of covert Scotticisms that retained their prevalence under the current of anglicisation. This complementary aspect to the study of anglicisation via usage guides can offer a more complete view of the diachronic development of Scots and Standard Scottish English forms and the relation between them. More will be said on this in Chapter 3.

It should be noted at this point that many of the forms referred to as Scotticisms by prescriptivists of the time are not necessarily exclusive to Scottish English, but also to some degree vernacular universals, i.e. found in many other non-standard varieties. An example is the use of *shall* vs *will* in the future tense of first and third person, which also shows up frequently as a feature of 'vulgar Irish' and a variety of other vernaculars. This is in fact something that is pointed out for a number of features in the usage guides collected below (note, for instance, Mitchell [1799: 78]'s following entry: "I *seed* him yesterday; Scotch, Irish, and vulg. English.—I *saw* him—"). Since this study is concerned with linguistic attitudes, hinging on what people at the time perceived as markedly Scottish, the forms will be referred to as Scotticisms throughout; however, it is good to keep in mind, heading into the remainder of this study, that this term does not always cover the full range of a language feature's usage.

As for the usage guides, a selection had to be made of works listing proscribed Scotticisms to use in the present study. Guided by Cruickshank (2011), a selection of four works was chosen, on the basis of which guides were the most influential during their own time but also the decades following their publication. As such, these usage guides likely had the largest reach and impact on the language of various Scottish social classes. Each of these works is outlined briefly below.

David Hume, *A list of Scotticisms* **(first published in 1752)**

David Hume was a well-recognised figure of the Scottish Enlightenment, having published a number of influential philosophical and historical works; his contribution to the landscape of linguistic prescriptivism is not as famous, but nonetheless one of the more influential contributions of the Scottish era of standardisation. Hume was highly concerned with ensuring that his own language consisted entirely of Standard English, and for that reason compiled a list of one hundred Scotticisms that included features that he had noticed both in his own language use and in that of the Scottish gentry with whom he interacted. This list was initially published in his work *Political*

Discourses (1752), and in following decades reproduced multiple times, in the Appendix of the 1760 Scots Magazine and again in the April 1764 volume of the same publication. The numerous republications suggest that the relevance and social value attached to Hume's list of Scotticisms remained strong throughout the decades, and the (largely upper-class) readership of this publication would benefit from these usage tips ongoingly.

The ordering of Hume's list was essentially non-existent. Dossena (2005) has grouped the items according to grammatical category, resulting in the following distribution: 33 verbs and verb phrases; 18 nouns and noun phrases; 14 prepositions and prepositional phrases; 12 adjectives and adjectival phrases; 12 adverbs and adverbial phrases; five entries on word order; four on pronouns; one on conjunctions; and one on idioms. Moreover, the list was preceded by a multi-paragraph discussion of the usage issue of *will* versus *shall*, which is a feature that shows up across practically all usage guides written during this period.

James Beattie, *Scoticisms: arranged in alphabetical order, designed to correct improprieties of speech and writing* (1787)

Hume's work may have been the first of its kind in applying prescriptivist ideas to Scots and Scottish English, but it was far from the last. Soon after, it was followed by the similarly influential *Scoticisms*, published in 1779 by James Beattie. Beattie, a philosopher by trade as well as a professor and a poet, found that his native Scots was "barren in itself" (cited in Cruickshank 2011: 22), and compiled a work detailing and proscribing various Scotticisms in that vein. Beattie's list, which is more extensive than Hume's at around 200 entries, shows a good deal of overlap with the latter; indeed, it includes over half of the entries listed by Hume. The Scotticisms are accompanied by a little more explanation of the forms and commentary to the Standard English alternatives offered than Hume gave, though such commentary is still few and far between. Beattie's features largely fit into the same linguistic mould that Hume's do, consisting of lexical, semantic, and morphosyntactic entries, as seen with the previous work. The 1787 republication of this guide, titled *Scoticisms: arranged in alphabetical order, designed to correct improprieties of speech and writing*, expands on the 1779 list with additional entries while also being more easily navigable due to its alphabetical ordering.

It is worth mentioning that Beattie remarks, in the 1787 republication of this guide, on his selection of Scotticisms to discuss: "With respect to broad Scotch words I do not think any caution

requisite, as they are easily known, and the necessity of avoiding them is obvious" (James Beattie 1787: 2-3). This suggests that an unknown number of the most common and overt Scotticisms are omitted from this guide; while Beattie evidently does not mention which features he has left unaddressed here, this means that some of the most societally marked Scotticisms are not being included here, something which may well go for the other usage guides too, though most make no specific mention of this. Some of the more overt Scotticisms that this study aims to investigate through the use of these usage guides, then, may paradoxically be going unobserved anyway because they were deemed 'too obvious' to include.

John Sinclair, *Observations on the Scottish Dialect* (1782)

Just a few years after Beattie's publication, the *Observations on the Scottish Dialect* (1782) by John Sinclair entered the market of Scottish prescriptivist works. Sinclair, a well-known and -respected member of the Scottish upper class, had a broad range of interests, which had him practising law, working as a Member of Parliament, and compiling various publications of the *Statistical Account of Scotland* (21 volumes, 1791-1799) on the Scottish population and parishes, which are still used for historical research today. His writings on Scots and Scotticisms are far from his best known, but *Observations* is nonetheless among the most influential of the usage guides published throughout this period. Cruickshank (2011: 25) suggests that nation-building was a strong motivation for him in urging his readership to speak one unified language in Standard English; Sinclair himself writes, later, that "[Observations] was drawn up, with a view of diminishing, as much as possible, the distinction between the two nations, in regard to language" (Sinclair 1831: xxviii).

One innovation in Sinclair's *Observations*, vis-a-vis the previously discussed guides, is his ordering of the entries. He does so by general topic, but also by grammatical category more specifically (Chapter 2, *Words peculiar to the Scots or, which they use in a sense different from the English*, is subdivided into verbs, adjectives, nouns, and particles). Specific domains that are covered in subsequent chapters include legal terms, as well as those pertaining to clothing, agriculture, furniture, etc. All in all, it covers some 900 features, including a large portion of the Scotticisms discussed by Hume and Beattie.

Hugh Mitchell, *Scotticisms, vulgar anglicisms, and grammatical improprieties corrected* **(1799)**

The final usage guide included in the present study is one that moves away from the previous ones in terms of the author's social background and intention in the writing of his usage guide. Hugh Mitchell, the son of a farmer, worked as a schoolmaster at initially the academy in Greenock, before moving to the English and French Academy in Glasgow. Cruickshank (2011: 30) writes that it was likely his teaching that motivated him to write his own usage guide, which as a result had a more overtly pedagogical purpose than the previous ones. Mitchell aimed his work particularly and explicitly at the urban middle classes, who, lacking in or in spite of their education, needed to master the societally advantageous English language, and avoid Scottish "colloquial words and phrases" (Mitchell 1799: vii). This, Mitchell believed, would allow them a path to societal advancement. Mitchell's list was thus aimed at a less upper-class audience than the works of his predecessors, such as Hume's, were; however, it stands out that even in his work, the lower classes, whose language is characterised according to Mitchell himself by "numberless uncouth Vulgarisms" (1799: vii), are not actually the intended readership, despite their language ostensibly being that which, according to the authors of these usage guides, was the most in want of their guidance. Mitchell's list, much like those of Beattie and Sinclair, is organised alphabetically, and touches on topics of grammar, etymology, as well as literary use of its entries.

An author not included here, though potentially interesting, is James Elphinston, who wrote *Propriety Ascertained in her Picture, Volumes I & II* (1786: 1787). While Elphinston wrote with the intended audience of a broader, less-extensively educated population than the previously mentioned authors, and as such might have been an interesting case study too, Beal (2003: 10-11) has ascertained that contemporary attitudes did not attest any great reach of the work, which "would be doomed to obscurity" (Cruickshank 2011: 27), largely because he insisted on using an idiosyncratic spelling throughout the work. As such, it was likely not influential enough, certainly over half a century later, to merit an individual investigation within the scope of this research.

3. Methodology

3.1 Building the corpus

In order to analyse the different degrees of anglicisation versus Scottish English in nineteenth-century Scottish correspondence, two datasets were compiled. These datasets, as discussed above, consisted respectively of archives of nineteenth century Scottish Pauper Petition Letters from the Tongue and Perth parishes, and correspondence taken from the Corpus of Modern Scottish Writing, or CMSW corpus. These datasets will be referred to as the ScotPP dataset and the CMSW dataset. The compilation and decision-making with regard to each of these datasets is outlined below.

3.2 The ScotPP dataset

For the ScotPP dataset, I drew on a research project conducted under the leadership of Dr. Moragh Gordon and Dr. Jelena Prokic at Leiden University (2022). This project sought to compile, transcribe, and digitise a large number of pauper petition letters collected from the parish archives of Tongue and Perth into a single corpus, titled the ScotPP corpus and totalling at 13.209 tokens. More on that corpus can be found on the corpus website, scotpp.lucdh.nl, but I will outline the basic steps of making the data available here.

First, we collected the letters from various Scottish archives. We ended up with a dataset of 13.209 tokens in total, from sources collected from two periods: a series of sources from the year 1821 from the parish of Perth, and sources from the parish of Tongue written between 1850 and 1852. It is important to note that the materials for both of these data subsets did not exclusively consist of pauper petition letters. Both datasets also included separate notes most likely written by their respective parish inspectors, cataloguing the letters with the added mention of whether the petitioners' requests had been granted or denied. Moreover, the Tongue dataset included a number of medical reports, written up by what appeared to be the local doctor in response to the petition letters in order to determine whether their medical situations were indeed as described by the petitioners. These reports, which were relatively short in length compared to the letters themselves, could also document home situations or any other features of note that had arisen during a visit; the reports were diplomatically transcribed and filed along with the PPLs. All the material included in the ScotPP, both petitions and additional documents, were used in the present study.

In total, the Tongue data included 54 petition letters; the dataset also featured nine medical reports, all by the same doctor R.W. Black, and 61 notes written by the inspector to catalogue the petition letters (some of these not being attached to any specific PPL). In total, this dataset consisted of 10.086 tokens. The Perth dataset, by contrast, was markedly smaller; it consisted of just 11 petition letters, many of which (though not all!) were, as previously mentioned, written by the same scribe, an Alexander Mackenzie. This dataset also included 13 cover notes—again, more than the number of petition letters overall, leaving some of the notes orphaned without their corresponding PPL. In addition, this dataset included two documents that catalogued the court proceedings of one Margaret Spence; these documents were transcribed and included, and appear to have been written by Margaret Spence herself. In total, the Perth dataset consisted of 3.123 tokens.

Overall, a diplomatic approach was chosen, marking a variety of features such as superscript items, margin-located notes, smudged writing, overwriting, and more.[1] Following the transcription stage, the letters were digitised, encoding them in a TEI format easily navigable and searchable for future researchers. Moreover, metadata was provided on the letters, including elements like time and place of writing. The digitisation, publication, and annotation of these materials allows researchers in fields such as historical (socio)linguistics and history to conduct further research into the compiled letters.

3.2.1 Metadata and sociohistorical features of note

A number of observations can be made regarding the social metadata of the authors. There was, with many of the letters, the question of authorship. While most PPLs appear to be written by the petitioners themselves, some letters were written by other scribes, likely often in situations where the petitioner themself was less educated, illiterate, or disabled. Identifying the letters that fell under this category was at times a delicate task. Many could be caught due to the inclusion of a 'her [x] mark' or 'his [x] mark' inserted into the signed name of the petitioner; the 'x', in these cases, served as the petitioner's signature in the document, which was for the rest written up by a scribe. However, the absence of this X-mark signature did not always unequivocally indicate that

---

[1] While the specifics of the transcription process regarding which features to describe and how to describe them are not discussed in this research, a more detailed overview of the transcription conventions followed for this corpus will be made available in the corpus manual, accessible via scotpp.lucdh.nl.

the PPL had been written by the petitioner themself, and careful consideration of the various handwritings found in the dataset was necessary in order to determine this. That said, the final conclusion reached on authorship could, for many of the letters, not be attested with one hundred percent certainty.

Since the question of who wrote a letter was fairly relevant for the purposes of this research, the letters included were tagged for having been written by the petitioner themself or by an external scribe. Moreover, a number of other features were also tagged within the dataset, wherever it was possible to determine them. These features included the gender of the writer, the town and/or region in which the letter was written, and the decade in which it was written. These factors, taken from Biber (1994)'s framework of situational factors that might influence variation in language use, were the ones that could be ascertained with reasonable certainty, considering the general lack of available metadata and in-depth information on the petitioners of these letters. The places in which they differed from the CMSW dataset, as discussed below, were interesting as potential explanations for disparities in the linguistic choices, particularly as relating to anglicisation and the inclusion of Scotticisms, between the two datasets.

3.3 The CMSW (middle-/upper-class) dataset

For the second dataset, which served to represent written correspondence from the same period as that of the ScotPP dataset but written by individuals of a higher social class, I compiled a selection of letters by making use of the Corpus of Modern Scottish Writing (1700-1945), created at the University of Glasgow by Wendy Anderson, Jenny Bann, and David Beavan (2007). This electronic corpus, henceforth referred to as the CMSW, compiles written text, including a large amount of correspondence, in a period that falls in between that covered by the Helsinki Corpus of Older Scots (covering the 1450-1700 period) and the Scottish Corpus of Text and Speech (1945-today). As such, it is highly useful for the present goal of comparing the language of early- to mid-nineteenth century PPLs to the language of their higher-class contemporaries in their correspondence.

The CMSW dataset was compiled by selecting written correspondence from the period of 1820 to 1860, as that is the period that was covered by the PPLs documented in the previous dataset. From the available materials of the CMSW, 44 were selected, on the basis of the listed metadata on the writers of the data. Occupation was taken into account here in particular, as the

most visible and reliable predictor of social class out of the metadata options offered in the corpus. As mentioned in Chapter 1, assigning social class to historical individuals is not without its pitfalls; this went all the more so for the CMSW authors, whose writing style and purpose did not allow itself to be categorised as clearly as the genre of petition letters did in the ScotPP dataset. Given the often limited amount of data available on the historical individuals whose language and writings are the subject of this study, a large factor in assigning people a particular class status for the purposes of characterising their data was their occupation, which information was usually available.

This dataset, at 14.742 tokens in total, was slightly larger than the PPL dataset. There were differences between the letter lengths, with the shortest letter included totalling just under 50 words; on the whole, though, the average number of words measured around 500. The writers of the letters had a number of different professions and geographical backgrounds; some, such as David Livingstone, were originally Scottish but wrote from entirely different continents, whilst others, such as James Hogg and Mary Mather, were situated in Scotland themselves when composing their letters. Some of the professions held by the authors included publisher, explorer, and author. The subjects and purposes of the letters covered both professional matters—reprints of published works, discussions of financial transactions, etc.—and informal ones, discussing personal matters such as new maids, appointments set for upcoming familial meetings, and more.

The same external factors, social and situational, that were tagged in the ScotPP dataset were tagged in these letters; that is, whether the letter was written by the petitioner themself; the gender of the writer; the age of the writer; the town/region where the letter was written; and the decade in which the letter was written. Another factor that was of interest was the relationship between addresser and addressee, which Biber (1994) mentioned as one of the unfixed variables that affect linguistic choices in writing. The social role relations in the PPLs were fixed, and unequal; the writer of the letter was by definition in a socially lower-ranking, subservient role to the addressee, hence the nature of the letter and the request. This difference in social position, something which might have a distinct effect on the linguistic activity, was generally not present in the letters of the CMSW dataset, where participants in interactions were of a similar social status, and familiar enough with one another that a level of distance and formality was removed from the interaction.

3.4 Selecting Scotticisms

As outlined in Chapter 2, the selection of variables to be investigated in the ScotPP and CMSW datasets was done on the basis of the most influential and widely known usage guides written by contemporaries around, or slightly prior to, the period in which the letters were written. These usage guides were the ones written by Hume (1752), Sinclair (1782), Beattie (1787), and Mitchell (1799). Of these usage guides, Mitchell's was the one that most clearly aimed itself at a readership of lower- to middle-class Scottish people, while the others focused more on prescribing the language of the upper classes.

Due to the large overlap in the language features these authors comment on, the lists could be easily combined into one extensive set of Scotticisms that were, according to the leading voices in Scottish prescriptivism of the day, the most notable (and most ill-advised) features of Scottish English. This list, which totalled at 526 features, included Scotticisms in a variety of linguistic domains. Most frequently, the proscribed Scotticisms were lexical, such as the following:

1. *Decreet*, for 'decree' (discussed by Beattie)
2. *Part with child*, for 'miscarry' (discussed by Hume)

Other types of Scotticisms also contained generalisable morphosyntactic rules, though these were not always listed by the authors with an awareness of that fact. To illustrate, Beattie discusses various instances of Scots speakers using perfect tense forms where preterite forms would be preferred in Standard English, and vice versa:

3. *Broke*, for 'broken'
   Proven, for 'proved' (both discussed by Beattie)

The Scotticisms were as such labelled according to different categories: those dealing with lexical items, those dealing with morphosyntactic items, and those dealing with orthographic items (particularly Sinclair focuses on the latter, with a separate section within his usage guide dedicated to proscribed spellings). The division of Scotticisms according to those categories was as follows:

- 466 lexical Scotticisms

- 41 morphosyntactic Scotticisms
- 19 orthographic Scotticisms

A full list of the Scotticisms taken from these usage guides can be found in Appendix 1.

3.4.1 Scotticisms not proscribed in usage guides

A second part of the analysis involved cataloguing the Scotticisms that occurred throughout the datasets which were not proscribed in usage guides of the time. These Scotticisms, which could give insight into the covert and overt Scotticisms that existed in the early to mid-nineteenth century, could of course not be selected on the basis of said usage guides. As such, a rather different source was consulted to aid in identifying these features in the data: the work of Dossena (2005), in which she compiles various lists of contemporary observed Scotticisms which were themselves compiled by linguists over the preceding decades. She includes the works of Aitken (1979, 1992), McClure (1994), Görlach (1990), Lass (1987), and Miller (1993), to create a chapter that showcases the various features that are considered defining ones of modern-day Scots by the mentioned scholars. In most of these authors' lists, features that are considered 'defining' include both covert and overt features, as well as the full range of possibilities that exist between them. Moreover, the Scotticisms vary by linguistic category: of the included authors, Miller (1993) is the one who most comprehensively touches on syntax, morphology, and discourse strategies, where the rest largely focus on lexical features—though items such as idioms also receive some attention. The full collection of Scotticisms can be found in Dossena (2005: 18-36).

3.5 Analysis of the data

The next step was then to investigate the two datasets compiled, the PPL and the CMSW data, for occurrences of the Scotticisms that had been discussed by the previously listed authors. The data was retrieved through a process of close reading the compiled letters and coding the relevant findings. In doing so, the entries taken from the usage guides were treated as variables—that is, linguistic items that could take the form of the Scottish English, proscribed form discussed in the usage guides, or the Standard English form that the authors prescribed. In this way, every occurrence of either variant throughout the letters could be annotated as either a Scots or a Standard English form. The total number of occurrences of these respective categories of variants were

compiled, and ratios of the occurrence of Scottish English variants versus Standard English ones were calculated. This was done separately for the ScotPP and the CMSW datasets, as well as more specifically the Perth and Tongue subsets of the PPL data, in order to account for regional variation. Moreover, the occurrences and ratios of the different linguistic types of Scotticisms—the lexical, morphosyntactic, and orthographic variables, as discussed previously—were also determined where numbers were large enough to allow for it. This made it possible to draw up a more detailed picture of the degree of anglicisation throughout the datasets.

A second step of the analysis phase was to identify the occurrences of Scotticisms that were not proscribed in the usage guides; that is, the covert Scotticisms, as opposed to the overt ones, that occurred throughout the data. Here, too, the total number of occurrences across the two datasets were written up, and the Scotticisms that were identified could be investigated more closely, based on linguistic type, category, and the context of their occurrence.

4. Results

4.1 ScotPP dataset

First, the ScotPP dataset, consisting of a total of 13.209 tokens, was investigated for the occurrence of Scotticisms proscribed by the usage guides.

In total, 38 instances of proscribed Scotticisms were found across this dataset, covering 24 usage guide entries. Of these Scotticisms, 29 were classified as lexical items. Examples include the following:

4. Likewise I wish to say that my wife would rather take to two **Bairns** (William McKenzie, 30 March 1852)

   Prescribed form: **children** (Sinclair 1782: 111)
5. which is **instructed** by the certificate produced (Alexander Mackenzie, 24 October 1821)

   Prescribed form: which is **proved** by … (Beattie 1787: 12)

Most of the time, the usage guide authors provide no further information as to why a specific lexical form is proscribed, beyond its being a Scotticism and thereby undesirable. For the morphosyntactic Scotticisms, brief context and/or explanation is sometimes provided, particularly by Sinclair (1782). Nine Scotticisms were classified as instances of Scotticisms in the morphosyntactic domain. These include the following:

6. also that one shilling has been **keept** off this month (Robert Newland, 9 March 1852)

   Prescribed form: **kept** (e.g. Sinclair 1782: 69)
7. Humbly Sheweth That the petitioner **haves** presented two former petitions (Isabella Reid, 4 November 1851)

   Prescribed form: **has** (e.g. Beattie 1787: 9)

Example 6 might also be interpreted as an example of an orthographic Scotticism; however, Sinclair (1782) himself lists it as an example of a "false formation in the Scottish dialect" (69), distinct from his earlier list of points in which "the Scotch and English dialects […] differ in orthography" (55). Where present, such classifications provided by the authors themselves were used to guide this study's classification of Scotticisms as well. Interesting to note here, too, is the

resemblance of example 7 to the Northern Subject Rule, a characteristic feature of not only Scottish English but other Northern varieties too; however, the NSR is usually more likely to occur with an adjacent pronoun, and likely not applicable here. More will be said on this rule, and why it likely is not what is occurring here, in Chapter 5.

No orthographic Scotticisms were found in the PPL data. On the other hand, this dataset contained a total of 77 items that represented the prescribed, Standard English form of a Scotticism as listed in one or multiple of the usage guides. These items covered a total of 42 usage guide entries. The vast majority of these items was classified as a lexical item, 73 in total. These include the following:

8. Therefore the Petitioner is **at present** going from house to house seeking shelter (Elen Sutherland, 24 December 1851)

   Proscribed Scotticism: … is **presently** (Sinclair 1782: 32)

9. I **hope that** the Honourable Members will be Pleased to consider that I have neither Father nor Mother to support me (Betsy Manson, 13 March 1851)

   Proscribed Scotticism: I **am hopeful that** … (Beattie 1787: 11)

One item was classified as a Standard English variant of an orthographic variable, namely the following:

10. … which prevents her from **lying** down at night (Scribe for Catharine McKay, 4 August 1851)

    Proscribed Scotticism: **laying** (Beattie 1787: 13)

As for Standard English morphosyntactic features, an important point to note is that the counting of these forms presented a challenge. Since the forms prescribed throughout the usage guides used here were overwhelmingly standard throughout both datasets, the ScotPP and the CMSW letters, they would skew the picture of ratios of Scotticisms and Scottish English forms as they occurred throughout the data. This applies in particular to forms such as the Standard English form 'he/she/it will', as opposed to 'shall'; and the SE form 'he has' as opposed to the Scotticism 'he have'. These forms were used so ubiquitously throughout the data that the decision was made

to simply consider their Scotticism counterparts on a case-by-case basis, so as not to skew the data presented here. This, of course, is a salient choice in and of itself, as it reflects a standardness of the Standard English form in the (written) language of all the here-represented social classes of the nineteenth century already, at least regarding these forms. In any case, the information listed in the tables above reflects this decision.

As such, three items were highlighted as examples of morphosyntactic prescribed Standard English forms, including the one seen in the example below:

11. … which can be **proved** by investigation (Scribe of Isabella Reid, 4 November 1851)
Proscribed Scotticism: **proven** (Beattie 1787: 17)

The above example pertains to a Scotticism that is mentioned by a number of the usage guide authors included in this study, and is described by them as a feature that is particularly common in Scots; moreover, it is a perfectly acceptable form in modern-day Scottish Standard English today. While the specific verb *proved* occurs across a number of usage guides, indicating a high degree of markedness in the contemporaneous metalinguistic discourse, it is one case of the often-recurring prescriptivism pertaining to the 'time past'; Sinclair (1782: 71) writes that "[t]he Scots are also apt to mutilate the termination of time past", dedicating an entire page to examples of this usage pattern. It thus stands out that no instances of the proscribed forms of past participle were found throughout the data, neither forms like the one listed in example 11 nor forms like the one in example 12, taken from Sinclair's own work:

12. *Where was he educate?*
Prescribed form: Where was he educated? (Sinclair 1782: 71)

An overview of the figures is represented in Table 1.

Table 1.

Scots and Standard English forms in the ScotPP data.

Scotticism type	Scots form	Standard English form	Total
lexical	29	73	102
morphosyntactic	9	3	12
orthographic	0	1	1
Total	38	77	115

Overall, these figures show that the ScotPP pauper petition letters contained around one Scotticism from the list proscribed in the usage guides for every two Standard English forms prescribed there; in short, a 1:2 ratio.

A closer look into some of the intra-writer variation found in highly represented writers—for instance, the aforementioned scribe Alexander Mackenzie—will be provided in Chapter 5. For now, the individual results of the two sets of pauper petitions compiled in the ScotPP data, the Tongue letters and the Perth letters respectively, will be briefly expanded upon.

4.1.1 Tongue data

The Tongue data consisted of 10.086 tokens in total. Within this data, a total of 20 instances were found of Scotticisms listed in the usage guides of Hume, Beattie, Sinclair, and Mitchell. Of those 20 instances, 14 of the items consisted of lexical Scotticisms; the remaining six were instances of morphosyntactic Scotticisms. Moreover, the Tongue letters contained 50 Standard English forms listed in the above usage guide. Of these forms, 47 were lexical Scotticisms; the remaining three were morphosyntactic in nature. These figures are detailed in Table 3.

Table 2.

Scots and Standard English forms in the Tongue PPLs.

Scotticism type	Scots form	Standard English form	Total
lexical	14	47	61
morphosyntactic	6	3	9

orthographic	0	0	0
Total	20	50	70

4.1.2 Perth data

The Perth data consisted of a total of 3.123 tokens. Within these letters, a total of 18 instances were found of Scotticisms listed in the usage guides of Hume, Beattie, Sinclair, and Mitchell. Of those 18 instances, there was one instance of a morphosyntactic Scotticism; the remainder were all lexical in nature. The Perth letters also contained 17 Standard English forms listed in the above usage guides; all of these forms were lexical Scotticisms. These figures are detailed in Table 3.

Table 3.

Scots and Standard English forms in the Perth PPLs.

Scotticism type	Scots form	Standard English form	Total
lexical	17	17	34
morphosyntactic	1	0	1
orthographic	0	0	0
Total	18	17	35

4.2 CMSW dataset

Next, the CMSW dataset, consisting of a total of 14.742 tokens, was investigated for the occurrence of Scotticisms proscribed by the usage guides.

In total, 14 instances of proscribed Scotticisms were found across this dataset, covering 13 usage guide entries. Of these Scotticisms, 12 were classified as lexical items. Examples include the following:

13. I **think shame** for having been so long in writing to you … (James Hogg, 24 March 1821)
    Prescribed form: I **am ashamed** … (Beattie 1787: 22)

14. I have therefore, to save time, **desired Mr Wardlaw to** send a copy of the lease … (Archibald Campbell, 17 April 1823)

    Prescribed form: I have **told him to** … (Beattie 1787: 20)

The remaining two Scotticisms were classified as instances of Scotticisms in the morphosyntactic domain. An example is the following:

15. Time flies faster in this capital **than ever I have found** it … (Burnes, 28 December 1834)

    Prescribed form: … than **I have ever found** it (Hume 1752 [unnumbered])

16. … at what period my [¿] as a judge will fall to be **stopt** (Archibald Campbell, 22 December 1823)

    Prescribed form: **stopped** (Sinclair 1782: 69)

Much like in the ScotPP dataset, none of the orthographic Scotticisms listed by the usage guide authors were found in the CMSW data.

On the other hand, this dataset contained a total of 59 items that represented the prescribed, Standard English form of a Scotticism as listed in one or multiple of the usage guides. These items covered a total of 42 usage guide entries. The vast majority of these items was classified as a lexical item, 49 in total. These include the following:

17. Constable is very **ill**. (James Hogg, 24 March 1821)

    Proscribed Scotticism: … is **badly/poorly** (Beattie 1787: 5)

18. For God's sake Mr Murray accept the **bill** (James Hogg, 11 June 1821)

    Proscribed Scotticism: the **account** (Sinclair 1782: 111)

A further six items were classified as orthographic Standard English variables, including the following:

19. I have [... gone] **likewise** to Oliver & Boyde (James Hogg, 11 June 1821)

    Proscribed Scotticism: **likeways** (Sinclair 1782: 57)

Moreover, four items were found to be examples of morphosyntactic prescribed Standard English forms, as seen below:

20. I fear they **must have been** lost (Mary Somerville, 7 March 1849)
    Proscribed Scotticism: they **have been** lost (Beattie 1787: 11)

These figures are represented in Table 4.

Table 4.

Scots and Standard English forms in the CMSW dataset.

Scotticism type	Scots form	Standard English form	Total
lexical	14	49	63
morphosyntactic	2	4	6
orthographic	0	6	6
Total	16	59	75

Overall, these figures show that the CMSW letters contained around one Scotticism from the list proscribed in the usage guides to every four Standard English forms prescribed there, or a 1:4 ratio.

4.3 Gender

As a point of metadata of note, the influence of gender was investigated for the respective ScotPP and CMSW datasets. In the first case, that of the ScotPP letters, this was complicated by the fact that a large number of the letters were penned by scribes or did not contain a sign-off at the end at all, making it difficult to determine what the gender of the author had been. Taking this into account, the following results were obtained:

Table 5.

Occurrences of Scotticisms versus Standard English forms among men and women in the ScotPP corpus.

Gender	Scots form	Standard English form	Total
Female	8	17	25
Male	15	28	43
Unknown	25	32	43

These figures show that little observable difference was found between the levels of anglicisation found in the writing of women compared to men in the ScotPP data. The ratio of Scotticisms to Standard English forms is minimally lower for women, but that difference is so minute as to be basically negligible.

For the CMSW dataset, the gender of all writers was known. There, the following results were found according to gender.

Table 6.

Occurrences of Scotticisms versus Standard English forms among men and women in the CMSW corpus.

Gender	Scots form	Standard English form	Total
Female	2	10	25
Male	14	49	43

Here, too, the results show small differences in the ratios of Scotticisms to Standard English forms according to gender—the letters written by women have slightly fewer Scotticisms to Standard English forms as compared to those written by men—but once again, the overall differences here are too small to assign much significance to.

4.4 Additional Scottish English forms not mentioned in usage guides

The second part of this study, which looks at covert Scotticisms which occur throughout the letters but are not mentioned in the usage guides of the time, takes as its basis the list of Scotticisms common in today's language, as compiled by Dossena (2005). This list totalled around 250 Scotticisms, and includes a good portion of lexical forms but also morphosyntactic features, such as the retention of *not* where Standard English would use the contraction *n't* (originally identified as a Scotticism by Miller [1993]).

The analysis of these results took a different form, as the Scotticisms included in Dossena's (2005) work were not given a corresponding Standard English form. This is because the linguists who compiled these forms were not working with prescriptive intent, unlike the authors of the usage guides. Additionally, these forms reflect Scotticisms that have persisted to become commonplace in modern-day Scots, and are therefore not presented in the context of Late Modern Scots' anglicisation, in conjunction with a competing Standard English form. This meant, however, that the analysis had to be organised differently. While various potential Standard English counterparts to these Scotticisms were searched for throughout the data (e.g. *late morning for forenoon, needs to be counted* or *is in need of counting for needs counting*) many did not occur at all throughout the two datasets. That could suggest that the Scottish forms were so ubiquitous that no Standard English forms were in use, but it is perhaps more likely that other Standard English forms were in popular use at the time than those proposed here. In order to avoid skewing the results due to this lack of fitting Standard English variants, I chose to contextualise the occurrence of these Scotticisms not by comparing their frequency to their Standard English counterparts, but rather by comparing the ratios of non-proscribed Scotticisms between the ScotPP and CMSW datasets to the ratios of proscribed Scotticisms between the ScotPP and CMSW datasets. In other words, if the number of non-proscribed Scotticisms in the ScotPP dataset and the CMSW dataset are more or less equal, while the proscribed Scotticisms occur twice as frequently in the ScotPP letters as in the CMSW ones, that is a difference that could be interpreted and from which conclusions could be drawn.

These non-proscribed Scotticisms occurred noticeably more sparsely throughout the datasets of both the ScotPP and the CMSW corpora as compared to the proscribed Scotticisms. The total number of these Scotticisms is listed in Table 7; a full list of these instances can be found in Appendix 4.

Table 7.

Non-proscribed Scotticisms occurring in the PPL dataset.

Scotticism	Scots occurrences
close	1
not	40
never (as past negation)	3
should	3
can	1
auld	3

Note: the use of *not* instead of *n't* occurred in all cases but two.

These results show that in total, six kinds of Scotticisms included in Dossena (2005) occur in the ScotPP dataset. Of these Scotticisms, the most common one is the retention of *not* in negation, as mentioned above. This form is used near-exclusively, with just two instances of the contracted *n't* form occurring. The different negatory forms are exemplified below:

21. if you **don't** comply to my necesities I will appeal to some other Churt (Widow John MacKay, 3 February 1852)
22. the Honourable members will be Pleased to consider such an object the Like of *hir* **is not** on the Parochiell List (William MacKenzie, 13 August 1852)

Moreover, a number of the occurrences of covert Scotticisms involved the use of non-Standard modals, in these cases using *should* where Standard English would require *ought*, and *can* where Standard English would use *may*. Examples of these occurrences are given below:

23. if his circumstances were faithfully & truly represented it **should** have been increased (George MacKay, 13 August 1852)
24. there **can** be no doubt,- that if the woman is dependent upon her own exertions for support. she is a fit object for relief (R.W. Black, 16 March 1852)

The results for the CMSW dataset are shown in Table 8.

Table 8.

Non-proscribed Scotticisms occurring in the CMSW dataset.

Scotticism	Scots occurrences
not	94
never (as past negation)	1
undertaken	4
forenoon	1
advocate	
should	1
want binding	6
needs no counting	1

Overall, there are eight kinds of Scotticisms from the list compiled by Dossena (2005) that occur throughout this data. The total number of occurrences is higher still, at 15 items throughout the letters—excluding the occurrence of *not* for *n't*, which skews the total number since it is used in every instance of negation. Here, too, the majority of occurrences is thereby reserved for the features relating to negation and to modals. The *n't* contraction does not occur once throughout the letters, with *not* being used exclusively in its stead. Moreover, the modal form *should* occurs in place of Standard English's *ought* frequently throughout the letters, at a total of seven instances, such as in the following:

25. I do not think Miſs Balfour **should** administer till the case is decided (Archibald Campbell, 4 May 1823)

5. Discussion

5.1 Proscribed Scotticisms in the ScotPP and CMSW corpora

The first observation that can be made from the data is that the total number of features found, both those listed in the contemporaneous usage guides and those included in Dossena (2005)'s compilation of Scottish features, is remarkably low. This goes not just for the Scotticisms themselves, but also for their Standard English counterpart variants. To identify just 38 proscribed Scotticisms throughout the PPL data, for instance, means that it is only around half the letters that include such Scotticisms. The CMSW letters only featured 14 proscribed Scotticisms over a total of 44 letters and 14.742 tokens. The Standard English variants, while occurring markedly more frequently—more on that further down—still totalled at just 77 and 60 occurrences respectively. This is a point that will be further discussed in section 5.3.

Overall, proscribed Scotticisms occurred less frequently than their Standard English counterparts in both the ScotPP and the CMSW datasets. However, the ratios of Scotticisms to Standard English variants in the respective datasets differed strongly. In the ScotPP dataset, the number of Scotticisms to Standard English forms was approximately twice as high as it was in the CMSW data, at 38 to 77 (or nearly 1:2) compared to 14 to 60 (or just under 1:4). These results show that the number of Scottish forms compared to the number of Standard English—that is, anglicised—variants is markedly lower in the writing of the early to mid-nineteenth century middle and upper classes than it is in that of the lower classes at the time.

It is important when considering these results to highlight that basing conclusions off the linguistically biased source of usage guides means that those conclusions will exist in the context of a specific metalinguistic perspective. The authors of usage guides select the features they include based not on large scale linguistic research, but on personal observations that are likely coloured by their own linguistic ideologies. As such, while the above results reflect the occurrence of Scotticisms that were seen as marked and stereotypical in the contemporaneous linguistic discourse, they are not necessarily reflective of the actual language use of the various social classes at their time of writing. Since those linguistic ideologies are also a large part of what motivated the process of anglicisation, usage guides can offer a valuable guide to those ideologies and how reflective they were of actual language use. However, at this stage, it is also important to keep in mind that these results should be read with this caveat in mind: while the design of this research is oriented towards anglicisation and the results can shed light on that ideology-driven process, they

should at this point not be taken as an unbiased, definitive account of the writing of the lower, or indeed upper, classes of the time.

It is important to keep in mind that the usage guides used here, which Cruickshank (2011) deemed among the most influential ones of the era, were published multiple decades prior to these letters' time of writing. The main target audience of these usage guides were middle to upper-class populations, whose Scotticisms were deemed something to erase from their speech in order for them to have the potential to rise to the ranks of English upper society. Usage guide author Sylvester Douglas, for instance, aimed in his Treatise to reach an audience "whose language has already been in a great degree refined from the provincial dross, by frequenting English company, and studying the great masters of the English tongue in their writings" (Douglas 1779, quoted in Jones 1991: 101). In fact, some of the authors even mentioned outright that they were disregarding the Scotticisms common to the lower classes, focusing entirely on those forms that their intended readership was suggested to be guilty of using; Sylvester Douglas announced that he was not concerned with educating those people who commit "the grosser barbarisms of the vulgar Scotch jargon" (Douglas 1779, quoted in Jones 1991: 101).

It is interesting, then, that by the time the mid-nineteenth century rolled around, the above results show that the Scotticisms listed in those older usage guides were found more frequently in the language of the lower classes than they were in that of the middle and upper classes. While these Scotticisms are suggested in the usage guides to be largely characteristic of the language of the Scottish higher classes, this had by then apparently shifted to be more true of the lower classes. These results seem to suggest that by that time, these forms had been largely prescribed out of the language of the higher classes; being the target audience of many of the usage guides, perhaps the prescriptivism those guides represent brought about a level of anglicisation in those higher classes that did not reach the lower classes.

Of course, it is possible and even likely that the usage guides alone were not the (sole) driving force behind these changes. The genre of the usage guide was only popularised around the mid-eighteenth century, while the anglicisation of Scots had been gradually growing for well over a century by that point. It is therefore important to acknowledge the option that the prescriptivists who authored the guides were not setting the trends in prescriptivism that are described here, but rather signalling pre-existing metalinguistic trends, and spreading and perpetuating the developments via their writing. What they prescribed may have been naturally and

sociolinguistically less frequent in these social classes to begin with. Regardless, though, the results here do suggest that there was a distinct development to speak of.

Comparing the rates of lexical to morphosyntactic and orthographic Scotticisms for the respective datasets yielded just minor results. In the ScotPP dataset, lexical Scotticisms made up 76,3 percent of the total number of Scotticisms. In the CMSW data, those lexical Scotticisms comprised 87,5 percent. The remainder was in both cases made up of morphosyntactic Scotticisms. These numbers suggest that the lower-class writers of the ScotPP letters were slightly more likely to use morphosyntactic Scotticisms in their writing than the higher-class writers, whose writing contained near-exclusively lexical Scotticisms. This is interesting because the use of morphosyntactic forms is usually less conscious than that of lexical forms; if the anglicisation of the higher classes was indeed further along than that of the lower classes, then it is unexpected that their language would retain more of the lexical Scots forms than the morphosyntactic ones. Overall, however, these differences were not big enough to draw any major conclusions from.

5.1.1 The ScotPP dataset

The pauper petitions of the ScotPP corpus included around twice as many Standard English variants as the number of proscribed Scotticisms. These numbers suggest that the Scotticisms that had been at the forefront of the metalinguistic discussion in late eighteenth-century Scotland were, by the early to mid-1800s, far from ubiquitous among the Scottish lower classes. Standard English forms, it seems, had crept a good distance into the language of these social groups by this time, after the de-anglicisation of the fifteenth and sixteenth centuries had given way to the anglicisation process of Scots.

The lack of available lower-class data from the time when the usage guides were written, the late eighteenth century, means that the diachronic process of anglicisation cannot be attested as accurately as one might like; this point will be elaborated on in section 5.3. Hence, it is difficult to attribute the high levels of Standard English in this data directly to the influence of usage guides. What the usage guides do offer here is an illustration of societal metalinguistic discourse and the direction of linguistic trends. In combination with a comparison to the number of Standard English forms in middle to upper-class Scottish writing during the time, a clear image arises regarding the degree to which anglicisation had progressed at that time in the language of the ScotPP corpus.

In any case, it seems that overall, the number of the 'vulgar' Scotticisms that contemporary commentators were so concerned about was, by the 1820-1850s, not as high in lower-class Scottish English as was suggested by many of the usage guide authors. For instance, the "numberless uncouth Vulgarisms which are peculiar to the lower class of people in Scotland" that Mitchell (1799: vii) considered all but beyond his help, seem rather outnumbered by the forms that he advocates for in his usage guide.

Noticeable right away is that the vast majority of the proscribed Scotticisms found throughout the ScotPP letters consists of lexical features. As mentioned, this occurred for both the PPLs and the CMSW letters, and is likely attributable to the fact that the usage guides themselves also contained a lot more lexical items than items of a morphosyntactic, orthographic, or discourse-organisational nature. As such, this finding makes sense given the nature of the data.

The number of morphosyntactically categorised Scotticisms was, by contrast, markedly low, with only a few entries in usage guides containing generalisable morphosyntactic features showing up throughout the data. One such feature, which did not occur at all in the CMSW letters, was the mismatch between first and third person pronouns and verb, as exemplified by the following:

26. That the Petitioner is in indigent circumstances **have** five of a weak family (Widow Robert Sutherland, undated [1851])

At first sight, this feature seems like it could be an example of the Northern Subject Rule (NSR), a subject-verb agreement pattern common to many varieties of English including Scottish and Northern Irish varieties (e.g. Pietsch 2005). However, the NSR usually triggers inflection when the verb is far removed from the subject, which is not the case in many of the examples of this form throughout the data, such as that in example 26. In this case, I believe that the NSR is not what we are seeing; the mismatch between first and third person verb conjugations seen in the above examples is one that is likely driven by external factors. One of the structural rules of pauper petitions was that they were required to be written in the third person form. Letters nearly always started with the phrase "The petition [...] humbly sheweth that the petitioner …", and the rest of the letter then fell under the subordinate clause of this phrase, with petitioners writing about themselves as 'the petitioner'. This was a structure that was not without its issues, and many

petitioners throughout the corpus evidently struggled with maintaining the appropriate form throughout their letter. Petitioners would at times briefly switch to using first person pronouns and verb endings, which looks like a Scotticisms as listed in the usage guides but may be more a feature of the petition genre than one of Scots. An example is the following:

27. That **your petitioner**, being confined & afflicted for three years by a painful disease, and one that has rendered **me** now quite helpless, **find** it altogether impossible to live on the hitherto small allowance **I have been** receiving. (Anne McKay, May 4 1851)

Additionally, something that stands out is the fact that the total section of the PPLs that offered space for naturally occurring Scotticisms was relatively low. This was because a large portion of the body of the letters was taken up by general formulaic language. Such formulaic phrases included the following:

28. The Petition of Chirsty McKay at Rhitongue.
    Humbly Sheweth That (Chirsty McKay, 12 July 1852)
29. & while I **shall Earnestly Pray** (Elssy McKay, 28 December 1851)
30. I **hope that the Honourable Board will be Pleased to consider my petition** (Isabella McKay, 14 January 1852)

The presence of these formulaic passages is no great surprise; as mentioned by Jones and King (2016), the petition letter that was the standard form of pauper request during the nineteenth century was a form of request that was heavily rule-constrained, consisting of formal and often near-archaic passages. An example is the way that each petition starts, with the phrase "The petition of [name, residence] Humbly sheweth". The spelling and morphology of the final word here, *sheweth*, reflects how the formulas used in these letters lean towards the archaic—something that is reflected in different letter genres, too, as pointed out by Pietsch (2015) in his study of Irish immigrant letters. As a result, these formulas can at times reflect older Scots forms, including Scotticisms proscribed in usage guides, which are through convention retained in places where otherwise Standard English forms might occur. This includes examples like 31, where the (not

exclusively, but markedly) Scots form of *will* over *shall* is used in the formulaic petition expression despite the fact that usage guide authors strongly condemned such forms in day to day usage.

31. And your Petitioner as in duty bound **shall** ever pray (Widow Donald McPherson, 17 January 1851)

    Proscribed Scotticism: **shall** used for third person future tense, as opposed to the preferred Standard English **will**

Pietsch (2015) suggests that archaic letter-writing formulas like the ones listed above could remain in use for decades after the forms they included went out of use in day to day speech, because the format of these letters was passed down across generations by individuals whose levels of education and literacy were relatively low. This would be true of the people writing the letters of the ScotPP, but in the case of this corpus, the existing rigid structures of the petition letter are likely of bigger importance still in the cementing of these archaic formulas. After all, these formulas are also used by scribes throughout the data, particularly in the Perth letters, and those scribes likely had higher levels of education than the average lower-class Scotsman. Moreover, they might have been more familiar with the rules of petition writing, perhaps having had access to the petition writing manuals by writers such as Angel Day, one of many (English) writers of manuals on how to compose and structure letters in general but also petitions in particular (Daybell 2012: 70).

Another important aspect of the genre of petition letters is the relationship between writer and addressee that is inherent to the register. The pauper petitions were the most common way of requesting financial aid during the early nineteenth century; however, this was not the only form these letters could take, just a very specific one, characterised by the rigidity of its form and structure. It was already noted that much of the space in each petition was taken up by highly formalised phrases, with little to no variation between them. These phrases, and the overall structures of these letters, have been characterised by Jones and King (2016) as "rigid" and "supplicatory". The genre of the petition letter was so well-established that deviating from it was in this period essentially unheard of, which shows clearly in the collected PPLs.

This supplicatory style is not just a matter of form, but one of social etiquette, too. The writers of PPLs were engaging in a formalised and structured form of begging, with the social

roles that entailed; their own social positions were lower than those of their addressee(s), and as such, they had to match their linguistic style to the communicative act. The formality and rule-governed nature of these letters went hand in hand with a need to 'talk up'; that is, to accommodate their own linguistic patterns to the situation and addressee at hand. In a socially unequal situation like that of petitioning, this may have involved speakers adopting, where they could, the prestige language forms of higher social ranks—which, in this case, had been moving towards Standard English for well over a century. In short, it is possible that the high degree of anglicisation in the letters was a conscious effort on the letter writers' parts to modify their language, in order to fit the rigid and entreating nature of the interaction.

5.1.1.1 Tongue letters

Overall, we have seen that the letters from the Tongue dataset, while including a number of proscribed Scotticisms, contained decidedly fewer Scots forms relative to the Perth letters. A potential explanation for this is the fact that the Perth petitions were written a full three decades before the Tongue petitions. It is indeed possible that during that three-decade period, the anglicisation of particularly the lower classes would have proceeded a great distance. Particularly during a time when the Scottish metalinguistic discussion was shifting more and more to focus on middle-class and, in some cases, lower-class populations, the effects of prescriptivism may well have been reaching the lower classes more and more, with change in their language use ensuing at an increasingly high speed.

The previously mentioned study by historians Jones and King (2016) may also shed light on how the occurrence of Scotticisms found in these letters is reflective the sociohistorical position of residents of Tongue, and specifically petitioners to the parish. As Jones and King note, the mid-nineteenth century point was a time when the social relations between petitioner and parish minister were still highly hierarchical, something that is reflected in and of itself in the rigid structure of the petition that was still being used around this time. Since a petition letter "necessarily conforms to an inflexible formula that emphasises the humility of the appellant and the generosity of those to whom the appeal is made" (Jones and King 2016: 94), that context demands a register that is highly accommodating and formalised, just like the contents of the letter itself establish a clear hierarchy between addresser and addressee. Those relationships, the authors

note, "were defined by strict formality and rigid social boundaries which precluded any close familiarity" (2016: 98).

These levels of anglicisation in written language may therefore not necessarily have been reflective of writers' day-to-day spoken language. Individuals being taught to write often picked the skill up as a domain largely separate from that of spoken language (cf. Fairman [2012] for a more thorough discussion of literacy as a sociolinguistic concept). As such, the language they were taught to use in correspondence, particularly the kind of heavily structured and formalised correspondence of the pauper petition letter, may have stood separate from how they used oral language, or even written language of a less rigidly structured register (Daybell 2012: 70). This means that the non-deliberately used Scotticisms, or Aitken (1979)'s *covert* Scotticisms, would have been less present in the carefully and formulaically drawn up letters of this corpus.

5.1.1.2 Perth letters

An important point to note when looking at the Perth pauper petitions is that, as mentioned, a large number of them—eight of the 13 letters included in the corpus—were written by a single scribe. This man, one Alexander Mackenzie, appears to have stepped in as a scribe for the petitioners who were not literate enough to compose their own petitions. Pinning down Mackenzie's background with any certainty is difficult, but he appears to have been a carpenter and a ship's joiner, who moved around Scotland frequently but settled in Perth for a prolonged period of time. Given his overrepresentation in the total Perth dataset, it is interesting to specifically highlight his language use throughout the letters.

Overall, Mackenzie's proportion of Scotticisms to Standard English forms is roughly the same as that of the overall ScotPP dataset including all authors. A total of four Scotticisms occurs throughout his writing, versus 9 Standard English forms. The Scotticisms are listed below:

32. but has **stopt** short &will pay no more (Alexander Mackenzie, 22 November 1821)

    Prescribed form: **stopped** (Sinclair 1782: 69)

33. and was in consequence confined **to her Bed** for several weeks (Alexander Mackenzie, 15 May 1821)

    Prescribed form: **to bed** (Beattie 1787: 15)

34. Peter Liddle shoemaker in Perth & **presently** Private in the Royal Perthshire Mekka (Alexander Mackenzie, 12 July 1821)

Prescribed form: **at present** (Sinclair 1782: 32)

35. which is **instructed** by the certificate produced (Alexander Mackenzie, 15 May 1821)

Prescribed form: **proved** (Beattie 1787: 17)

With roughly twice as many Standard English forms as Scotticisms, Mackenzie's results correspond more with those of the wider ScotPP corpus (where the Standard English forms outnumbered the Scotticisms two to one) than with those of the remaining Perth letters (which had a one to one ratio). While the results remain so small in number that it is difficult to make any definitive generalisable statements, they do seem to show that Mackenzie, as the person doing much of the petition writing for the parish, was relatively aware of the prescriptivist language norms that rejected overt Scottish forms. The fact that he used fewer overt Scotticisms on the whole, at least in written language, indicates that the degree of anglicisation in Mackenzie's language was higher than that of his fellow parishioners who wrote the remaining letters in the Perth dataset. Moreover, it is likely that the further anglicisation of his language was causally linked to his suitability as a scribe for these parishioners. That suitability may have had several reasons: perhaps he was better schooled, more acquainted with the register in which pauper petitions were meant to be written, or some combination of these factors. In any case, his command of a more anglicised language than the other writers may have played a large role in making him suitable as a scribe. Ultimately, though, this does remain a matter of speculation.

5.1.2 The CMSW dataset

The results of the CMSW dataset are in many ways comparable to those of the ScotPP data. Here, the total numbers of both Scotticisms and Standard English features listed in the usage guides are also remarkably low, with only 59 instances occurring across a total of around 15.000 tokens. While the proportion of Scotticisms to Standard English forms is notably different for the CMSW letters, with one Scotticism for every four Standard English forms, these numbers are remarkably low, considering the number of authors publishing usage guides at the time of the letters' writing with the express intent to eradicate these forms. This is particularly true given that the writers of the CMSW letters were the people who formed the actual target audience of these usage guides.

The structure of these letters allows for a good deal more freedom in word choice; a much smaller portion of the CMSW letters is taken up by the formulaic phrases that were so common in the ScotPP data. While certain set phrases did still show up across letters—for instance, a variety of formulaic closing sign-offs, such as *I remain your Obedient and Faithfull Servant* or *Yours most truly*, are used throughout the data—the free body of the letters covers a much larger proportion of the individual CMSW letters than it did in the ScotPP petitions. This makes it all the more notable that the number of results for these letters was so low. The normalised occurrence of both proscribed Scotticisms and prescribed Standard English forms is thus even less frequent here when compared to the ScotPP results than it appears at first sight.

This infrequent occurrence of Scotticisms is, again, in part due to the specificity of the prescribed forms that occurred in the usage guides. Since those were frequently specific to specialised lexical domains, those items were less likely to occur in letters not explicitly related to those domains. Indeed, the vast majority of the Scotticisms observed in the CMSW data were of more general use than those lexically specific items that make up the majority of many of the usage guides. This means that examples such as 36 and 37, including more generalisable features, were more frequent than the more niche word in example 38:

36. he would require to see the warrant in order to ascertain at what period my [¿] as a judge will fall to be **stopt** (Archibald Campbell, 22 December 1823)
    Prescribed form: **stopped**
37. therefore let it be in a kind letter else I **will** be exceedingly grieved (James Hogg, 5 May 1821)
    Prescribed form: **shall**
38. he makes an assertion that the **stipend** is so much but leaves out of view the etceteras which make it much more (David Livingstone, 5 October 1858)
    Prescribed form: **salary**

Another interesting difference between this dataset and the petitions of the ScotPP corpus lies in the relationships between their writers and addressees. As noted in Chapter 2, Biber (1994) describes how this relationship can be influential in the selection of register and linguistic variables. As we have seen, the ScotPP letters were largely concerned with making requests, which

would place writer and addressee at different hierarchical levels, with the writer 'talking up' to the recipient from whom they expect or hope to gain something. This power imbalance would likely have had a significant influence on the formality of the language used in the writing. The letters of the CMSW corpus were generally not engaged in making requests (though it does occur in some instances, such as in Hogg [24 March 1821]). Instead, many of the CMSW letters were written between business partners, or between personal friends and family members. There are, for instance, the letters of Mary Mather to her sister Elizabeth Buchanan; the relationship between these women is naturally both more equal and more informal than that between the petitioners and inspectors in the case of the PPLs. A letter by Grace Knox, too, showcases a level of familiarity that could never have occurred in the pauper petitions of the ScotPP data:

39. I ought to have given it before parting with you but as taking farewell of my friend is always painful and confusing to me it quite escaped me till after you were gone. (Grace Knox, undated [1821])

This different, more balanced relationship between writer and addressee in the CMSW letters is particularly interesting because it might suggest that the language used throughout the letters would consequently be less formal, too. The fact that despite these more familiar dynamics between writers and addressees, the letters still contain such a low proportion of Scotticisms compared to the ScotPP letters, further strengthens that difference. Even in the distinctly more informal social context of the CMSW letters, the language of these middle and upper-class writers is apparently more devoid of Scotticisms than that of the lower-class writers writing in the highly formalised context of their pauper petitions.

It is worth noting here that while the writers of the CMSW letters included few Scotticisms throughout their written work, it is not necessarily the case that this translates to their spoken language also being near-entirely free of those Scotticisms. Görlach (1999: 149-150) has pointed out that letters "reflect the social and functional relations between sender and addressee to a very high degree—only spoken texts can equal this range." While he acknowledges that "[p]rivate letters can contain valuable evidence on informal usage", Görlach asserts that "[t]hey rarely include dialect […]. Writing is so much connected with the school and standard language that composing a letter in dialect is a breach of sociolinguistic convention". It is likely that the

education of the middle- and upper-class writers of the CMSW letters had, through their formal schooling, granted them a knowledge of the norms and formalities of written language that precluded any overt Scotticisms from entering into their writing. The next section will discuss this possibility, and how it ties into the presence of *covert* Scotticisms in those letters.

5.2. Non-proscribed Scotticisms in the ScotPP and CMSW corpora

There are a number of alternate explanations as to why the presence of Scotticisms throughout the ScotPP and the CMSW corpora was so distinctly lower than that of Standard English forms. One possibility is that the occurrence of colloquial features—for that is what Scotticisms had, in the nineteenth century, become—was less widespread in written language than it was in spoken language. This is, of course, a common challenge for historical sociolinguists looking to study spontaneous, oral-like language use: written language, even when in the form of ego-documents, does not typically reflect spoken language with complete accuracy. Since it is acquired differently than spoken language, in a classroom setting with an emphasis on formally dictated grammar norms, written language is often governed by a set of rules that prevent, in many cases, colloquialisms from entering texts in places where they would feature more extensively in spoken language. This higher level of conscious language use might have led Scottish people to phase out the elements of their language that were markedly Scottish to them (following Myers-Scotton [1998]'s Markedness Model), while retaining the features that were not so stereotypically Scottish as to be proscribed in the metalinguistic discourse, even though they were still distinctly Scottish. This effect was possibly strengthened by Scots and Standard English being far enough apart that using Scots might have felt to Scottish people like some kind of deliberate code-switching (for a more extensive perspective on historical code-switching and language attitudes in Scotland, see Dossena [2002]).

Here, again, Aitken (1979)'s terms *covert* and *overt* Scotticisms come into play, with covert Scotticisms being those used unselfconsciously by speakers/writers, and overt Scotticisms being tied to a higher level of awareness of their Scottishness—they are, in short, highly marked as Scottish, and thus proscribed more strongly. As such, it might be expected that if speakers have a higher level of linguistic (self-)consciousness (for instance on account of societal discourse condemning certain features as 'vulgar' and 'uncouth'), the features whose Scottishness goes unnoticed may be those that remain untouched, and persist in the language. By contrast, the overt,

markedly Scottish features—marked because, for instance, usage guides point them out as such and prescribe other forms in their place—would be the first to go in a process of anglicisation.

The remaining part of this analysis will focus on the characteristically Scottish features that occur throughout the data which are not listed in any of the usage guides as proscribed Scotticisms. These features, taken from the collection compiled by Dossena (2005), certainly occurred less frequently throughout both datasets than those proscribed by the usage guides of the time. There was, moreover, no strong difference between the frequencies of occurrence throughout the respective ScotPP and CMSW datasets—the results showed that seven of the features included in Dossena's work occurred in the ScotPP dataset, and eight of them in the CMSW dataset, numbers that, particularly if normalised, are basically equal (and equally low).

Non-proscribed Scotticisms that occurred in the ScotPP dataset included:

40. as I am an **auld** frail woman upwards of 80- years (Isabella McKay, January 14 1852)
41. which your Petitioner knows nothing of as he **never** at any time said to her that such a sum was due till he has given her a Charge to appear before your Lordship (Margaret Spence, January 13 1821)

The CMSW letters, moreover, included features such as:

42. if you have nothing particular to buy, come with the **forenoon** omnibus if in time (Mary Mather, April 29 1841)
43. The last edition of the Wake **needs no counting** for that was stipulated at £100. (James Hogg, March 24 1821)

The fact that the features listed by Dossena (2005) did not occur as often as those in the usage guides is interesting, because of how they fall on the scale of covertness to overtness as compared to the proscribed Scotticisms. The Scotticisms discussed here are the ones that were not deemed socially marked as Scottish, or at least not enough to be eradicated via usage guides. It is of course uncertain why these features were not stigmatised in the way that other Scotticisms were, left unacknowledged in a way that suggests that either these features were not recognised as Scotticisms or were considered less 'vulgar' or 'colloquial' than the previous features. It is likely

not their linguistic categorisation; these features, too, contain primarily lexical forms, though morphosyntactic features are also discussed. Regardless, it can be argued that it is that very lack of ostracisation that has allowed them to persist in present-day Scottish English. Where usage guide authors of the 1700-1800s did not consider these specific features in need of ostracising, today they are considered by linguists like Aitken (1992) and McClure (1994) to be among the strongest markers of the Scottish English language variety.

The most prominent feature in this section of the analysis is the use of *not* as opposed to a contraction *n't*. This feature occurred near-exclusively throughout this data, which was particularly notable because the contracted form had already entered into usage generally in English in the seventeenth century (González 2007); the near-complete absence of this contraction from the data therefore stands out. The contraction *n't* occurs twice in the PPLs, both times in the form of *don't*, as seen below:

44. if you **don't** comply to my necesities I will appeal to some other Churt (Anne McKay, February 3 1852)
45. I hope that the Honourable members will be pleased to consider that I **don't** wish to be placed upon the poors' List (Betsy Manson, March 13 1851)

For the rest, however, the contracted *n't* form does not show up at all, and it is entirely absent from the CMSW data. Instead, writers use forms such as the following:

46. I am happy to say we **have not** had much to complain of this year (Mary Somerville, April 3 1842)
47. but was last year deprived of the "land" as I **could not** pay any "Rent" (Angus McKay McNeill, June 16 1852)

Furthermore, it is interesting that in comparing the frequencies of these covert Scotticisms in the PPL and the CMSW datasets, the overall numbers of occurrences are more or less equal. With seven types of Scotticisms occurring in the ScotPP data, and eight types in the CMSW data, the numbers skew just slightly in the CMSW data's favour, though those figures are so small that it is a stretch to make much of that slightly larger number. The total number of tokens of those

Scotticisms is a little less equal for the two datasets, with the CMSW data slightly outweighing the ScotPP data. In total, non-proscribed Scotticisms occurred 11 times in the ScotPP data, excluding the near-universal occurrence of *not* over *n't*, the inclusion of which would cloud the results; for the CMSW dataset, that number is 15, again excluding uncontracted *not.*

This is noteworthy because, as seen above, the number of overt Scotticisms was markedly lower in the CMSW data than in the ScotPP data. These results can be read from two perspectives, which highlight different sides of the comparison: either the language of the lower classes features a markedly low proportion of covert Scotticisms in relation to overt Scotticisms, or the higher classes use a markedly high number of covert Scotticisms in relation to overt Scotticisms. The question is whether metalinguistic influences on lower-class language, middle-class language use, or of course a combination of the two, can explain these findings.

A potential argument for the latter is the level of metalinguistic awareness that middle- and upper-class Scottish people had access to during this time. As has been suggested in the discussion of proscribed Scotticisms, these social classes likely had a high level of awareness of the contemporaneous metalinguistic discourse, since prescriptivism had been widespread throughout their social circles for over a century by this time. While this metalinguistic awareness might have led them to use fewer markedly Scottish features in their language—those features, in short, that were proscribed in the usage guides of the time—the unmarked Scottish features would have been able to persist in their language use more freely without being suppressed with the goal of deliberate anglicisation.

The language of the lower classes, by contrast, was less affected by this suppression of markedly Scottish features, as we have seen above. Whether this had to do with a lack of immediate potential to move upwards in society through the adoption of anglicised speech and cultural forms, or simply a lack of awareness of the prescriptivist ideas that had propelled that anglicisation in the ranks of the middle and upper classes, the proportion of overt Scotticisms to Standard English forms in their written language was comparatively higher. This logically suggests, then, that the differences between the frequencies of overt and covert Scotticisms in the language of the lower classes would be less distinct than those of the middle classes. This is reflected in the fact that covert Scotticisms, in the ScotPP dataset, do not occur twice as often as they do in the CMSW data, as in the case of overt Scotticisms.

Overall, it is important to include this additional, non-prescription-based perspective in the discussion of Scots anglicisation because of the risks that studying a standardisation process brings with it. As mentioned, historical sociolinguistics' traditional focus on the language of upper classes and the standardisation thereof runs the risk of taking a teleological view on this process, thinking of Scottish Standard English as the end state and 'goal' of the linguistic development of Scots. Examining, too, the elements of Scots that were not anglicised, but persisted in the present-day situation of Scots, shows that the anglicisation process—a case of engineered standardisation from above—does not reflect the full picture of the development of Scots. The linguistic features that developed further and took root as established Scots forms, even as more marked features were being stigmatised in the metalinguistic discussion of the time, reflect a side of the linguistic development that is necessary in order to get the full picture of what was taking place in the linguistic landscape of Scotland at the time. This is helpful in stepping away from the traditional teleological perspective, and highlights the *language from below* approach that this study has sought to take.

5.3. Limitations, implications, and new insights

As was mentioned at the start of the discussion, the number of results found in the data is remarkably low, for both the ScotPP and the CMSW materials. This applies to the numbers of both Scotticisms and Standard English forms, and may have a number of potential reasons.

One factor is that the total number of distinct entries listed in the usage guides was simply not that high. The total number of entries taken from the usage guides, with entries that occurred multiple times across the guides collapsed into a single item, results in a total of 526 features that could be searched for throughout the two datasets—something that resulted in part, of course, from authors borrowing extensively from each other's lists of proscribed Scotticisms in compiling their own. Moreover, the type of Scotticisms that occurred most frequently across the usage guides also plays a strong role here: the vast majority of the Scotticisms included by the four authors consisted of lexical items, as opposed to more generalisable linguistic categories such as morphology, syntax, or even orthography. These lexical items would naturally be expected to occur less frequently than more morphosyntactically oriented, and therefore more productive, Scotticisms.

Moreover, in the case of the ScotPP dataset, the specific kinds of lexical items listed in the usage guides were not always from domains likely to occur often in the petition letter genre. The

domains covered and even categorised in the work of Sinclair were not necessarily exclusive to the language of middle and upper classes; universally common themes, such as food, or timekeeping, were represented alongside the more specific domains such as legal and agricultural terms. However, the specificity of these lexical entries meant that they would not occur frequently throughout pauper petition letters, which were usually concerned with general descriptions of poverty, illness and the like. It is noticeable, too, that as soon as letters were more concerned with legal proceedings—the Perth letters, for instance, include accounts by a Margaret Spence about the lawsuit being organised against her, and other letters refer to legal affairs the petitioners are involved in, too—the number of both Scotticisms and Standard English variants quickly rises, as the topic of the letter corresponds with a specific subset of Scotticisms, in this case one covered by Sinclair's *Observations*.

Unfortunately, the bad data problem remains a fact of life when conducting historical sociolinguistic research. While the present study has worked hard to minimise the disparities between datasets and create an optimal comparative environment, the limitations of the available data and methodology are present enough to require some qualifying statements.

Anglicisation is, by nature, a diachronic process. For the Scots language to go from the peak of its status to a low-prestige language variety condemned by commentators nationwide required time and gradual development. For this reason, the level of anglicisation in a given text would ideally be measured against a socially and contextually comparable dataset from before the material of interest, so that the diachronic difference might be attested. That data is, for the language of lower classes in Scotland, not available as of yet. The alternative approach, of measuring the degree of Scotticisms versus Standard English forms that have been made metalinguistically salient through their inclusion in usage guides, allowed for a perspective on the degree of anglicisation that still offers valuable insights into the linguistic situation of the lower classes of nineteenth century Scotland.

There is potential to build on these results in future research, however. Since there are likely a good deal more pauper petitions strewn across the depths of Scottish parish archives—some possible sources for the material are listed in Jones and King (2016: 86)—there is certainly potential for future studies to uncover more data and complement this work with that missing diachronic angle from which it would benefit. Indeed, the ScotPP is at present being expanded to include further Scottish PPL materials, thereby already broadening the range of materials from this

period available for research purposes. Bringing earlier data into the comparative mix might serve as an important corroboration of the results of the present study, which have shown the anglicisation of lower-class Scottish English to have been well advanced already by the early nineteenth century, though still distinctly less advanced than that of the middle and upper-class Scottish writers of the same time.

There are, moreover, further angles in the study of these pauper petitions that remain open avenues for further research; social factors such as gender, age, and education level may be of interest to both historical sociolinguists and historians, if such information can be identified reliably through further archival research. The pauper petitions represented in this study form a rich resource for such studies, and a resource that can be complemented with further materials in the future for a fuller picture of the linguistic and historical situation of nineteenth century Scotland.

Auer et al. (2015: 6) note that "[o]ne of the core concerns of historical sociolinguistics [...] is the effort to overcome the social bias connected to class, education and literacy inherent in written sources that has afflicted historiography". With the focus, particularly in the field of historical sociolinguistics, having shifted towards the *language from below* approach in recent research, the ScotPP materials provide a necessary and important perspective on those previously understudied language forms.

6. Conclusion

This thesis examined new data from lower-class Scottish writers in the early to mid-nineteenth century, making use of the newly assembled ScotPP corpus. These new materials represent an important step in the study 'from below' of historical Scottish English, and a fresh insight into the language of a previously understudied social group.

Comparing the ScotPP materials to a set of correspondence data from the same period written by middle to upper-class Scottish people, taken from the Corpus of Modern Scottish Writing, this study investigated the degree to which anglicisation had progressed in the language of the different social layers of Scottish society. The first section of the analysis, focusing on what Aitken (1979) termed *overt* Scotticisms, found that the Scottish features explicitly proscribed in the metalinguistic discourse of the era were markedly less likely to occur in the writing of the middle and upper-class writers' letters than in those of the lower classes. This difference suggests that while the prescriptivist ideals had permeated the writing of the higher classes of Scottish society, leading to the forfeiture of Scottish features in favour of their more prestigious Standard English counterparts, that process had not taken place to the same extent in the writing of the lower classes. However, those lower classes did include around twice as many Scotticisms as Standard English variants in their writing, suggesting that the beginnings of a process of anglicisation had nonetheless taken root in their language use, too. Further research, incorporating potential sources from the Scottish lower classes from a period prior to the one investigated here—if those sources can be uncovered—may shed further light on this development.

Additionally, the occurrence of covert Scotticisms was investigated, making use of a list of Scotticisms provided by Dossena (2005). Here, in contrast to the previous case, the frequencies of covert Scotticisms were found to be more or less equal across the ScotPP and CMSW materials, even skewing slightly in favour of the CMSW letters; in other words, it seems that the higher-class writers of the nineteenth century used more covert Scotticisms in relation to overt Scotticisms than the lower-class writers did. This further suggests that the awareness of higher-class writers that the proscribed Scotticisms should not be used informed their language use, while the covert Scotticisms, not attached to such a conscious metalinguistic attitude, were not particularly affected by the anglicising trend and writers' self-correction. By contrast, the language of lower-class writers was likely less governed by self-correction and the same levels of conscious metalinguistic

awareness. As such, their use of covert Scotticisms did not differ notably from that of the middle and upper classes of their time.

Overall, this study has opened new avenues for research into lower-class Scottish, in line with the desires within the field of historical sociolinguistics to no longer focus primarily on the standardised and upper-class forms that have traditionally been the focus. With a number of potential areas for future research suggested, there is a promising road ahead for developing a greater understanding of the Scottish *language from below*.

Bibliography

Aitken, A. J. (1979). Scottish speech: A historical view with special reference to the Standard English of Scotland. *Languages of Scotland*, *4*, 85-120.

Aitken, A. J. (1992). Scottish English. In T. McArthur (Ed.). *The Oxford Companion to the English Language* (pp. 903-905). Oxford: Oxford University Press.

Anderson, R. (1985). Secondary schools and Scottish society in the nineteenth century. *Past & Present*, (109), 176-203.

Anderson, W., Bann, J. & Beavan, D. (2009). Corpus of Modern Scottish Writing (CMSW). *Corpus Linguistics* 2009, 20-23 July 2009, Liverpool.

Auer, A., & Fairman, T. (2013). Letters of artisans and the labouring poor (England, c. 1750–1835). In P. Durrell, M. Scheible, S. Whitt, & R. J. Bennett (Eds.). *New methods in historical corpora* (pp. 77-91). Tübingen: Narr Francke Attempto Verlag.

Auer, A., Peersman, C., Pickl, S., Rutten, G., & Vosters, R. (2015). Historical sociolinguistics: the field and its future. *Journal of historical sociolinguistics*, *1*(1), 1-12.

Bald, M. A. (1927). The pioneers of Anglicised speech in Scotland. *The Scottish Historical Review*, *24*(95), 179-193.

Barclay, K. (2017). Negotiating independence: manliness and begging letters in late eighteenth and early nineteenth-century Scotland. In L. Abrams (Ed.). *Nine centuries of man: manhood and masculinity in Scottish history* (pp. 142-159). Edinburgh: Edinburgh University Press.

Beal, J. C. (2002). 'Out in Left Field': Spelling Reformers of the Eighteenth Century. *Transactions of the Philological Society*, *100*(1), 5-23.

Beal, J. C. (2004). *English in Modern Times: 1700-1945*. London: Routledge.

Beattie, J. (1779). *A list of two hundred scotticisms*. Aberdeen.

Beattie, J. (1787). *Scoticisms: arranged in alphabetical order, designed to correct improprieties of speech and writing*. Edinburgh and London.

Biber, D. (1994). An analytical framework for register studies. In D. Biber & E. Finegan (Eds.). *Sociolinguistic perspectives on register* (pp. 31-56). Oxford: Oxford University Press.

Corbett, J. (2013). The Spelling Practices of Allan Ramsay and Robert Burns. In W. Anderson (Ed.). *Language in Scotland* (pp. 65-90). Leiden: Brill.

Cruickshank, J. (2011). *The emergence of Scottish Standard English: the evidence of the correspondence of 2nd Earl Fife 1764-1789*. Doctoral dissertation, University of Aberdeen.

Daybell, J. (2012). *The material letter in early modern England: manuscript letters and the culture and practices of letter-writing, 1512-1635*. New York: Springer.

Dekker, Rudolf. 2002. *Egodocuments and history: autobiographical writing in its social context since the Middle Ages*. Hilversum: Verloren.

Devitt, A. J. (1989). *Standardizing written English: Diffusion in the case of Scotland, 1520-1659*. Cambridge: Cambridge University Press.

Dossena, M. (2002). A strong Scots accent of the mind: The pragmatic value of code-switching between English and Scots in private correspondence–A historical overview. *Linguistica e Filologia, 14,* 103-127.

Dossena, M. (2005). *Scotticisms in Grammar and Vocabulary. "Like Runes upon a Standin' Stane?"*. Edinburgh: John Donald.

Dossena, M. (2013). Ego Documents in Scottish Corpora: The Contribution of Nineteenth-century Letters and Diaries to the Study of Language History. In W. Anderson (Ed.). *Language in Scotland* (pp. 91-111). Leiden: Brill.

Elspaß, St., Langer, N., Scharloth, J. & Vandenbussche, W. (2007). *Germanic Language Histories 'from below' (1700–2000)*. Berlin: De Gruyter.

Fairman, T. (2012). Letters in mechanically-schooled language. In M. Dossena & G. Del Lungo Camiciotti (Eds.). *Letter writing in Late Modern Europe* (pp. 205-228). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Frank, T. (1994). Language standardization in eighteenth-century Scotland. In D. Stein & I. Tieken-Boon van Ostade (Eds.). *Towards a standard English*, *1600-1800* (pp. 1-62). Berlin: Mouton de Gruyter.

Geddes, A. (1792). Three Scottish poems, with a previous dissertation on the Scoto-Saxon dialect. *Archaeologia scotica*, *1*, 402-440.

González, M. P. C. (2007). *Uncontracted Negatives and Negative Contractions in Contemporary English: A Corpus-based Study*. Doctoral dissertation, University of Santiago de Compostela.

Görlach, M. (1990). *Studies in the History of the English Language*. Heidelberg: Carl Winter Universitätsverlag.

Görlach, M. (1999). *English in Nineteenth-Century England - An Introduction*. Cambridge: Cambridge University Press.

Houston, R. (1985). *Scottish Literacy and the Scottish Identity: Illiteracy and Society in Scotland and Northern England, 1600–1800*. Cambridge: Cambridge University Press.

Hume, D. (1752). Scotticisms. *Political Discourses* (unnumbered pages). Edinburgh: R. Fleming.

Jones, C. (ed.) (1991). *A Treatise on the Provincial Dialect of Scotland, by Sylvester Douglas.* Edinburgh: Edinburgh University Press.

Jones, P., & King, S. (2016). Voices from the far north: pauper letters and the provision of welfare in Sutherland, 1845–1900. *Journal of British Studies*, *55*(1), 76-98.

Joseph, J. E. (1987). *Eloquence and Power*. London: Frances Pinter.

Koch, P., & Oesterreicher, W. (1985). Sprache der Nähe—Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch*, *36*(1), 15-43.

Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Labov, W. (1994). *Principles of Linguistic Change. Vol. 1, Internal factors*. Oxford: Blackwell.

Labov, W. (2001). *Principles of Linguistic Change. Vol. 2, Social Factors*. London: Blackwell.

Lass, R. (1987). *The shape of English: Structure and history*. London: JM Dent & Sons.

Lowth, R. (1762). *A short introduction to English grammar*. London. Facs. repr. 1967. Menston: The
Scolar Press.

McClure, J. D. (1994). English in Scotland. In R. Burchfield (Ed.). *The Cambridge History of the English Language. Vol. 5, English in Britain and Overseas: Origins and Development* (pp. 23-93). Cambridge: Cambridge University Press.

Meurman-Solin, A. (1993). The author addressee relationship and the marking of stance in the characterisation of sixteenth and seventeenth century genre styles. *New Horizons in Stylistics? Belgisch Tijdschrift voor Filologie en Geschiedenis, 71*:3, 733-45.

Meurman-Solin, A. (1997). Differentiation and standardisation in Early Scots. In C. Jones (Ed.). *The Edinburgh history of the Scots language* (pp. 3-23). Edinburgh: Edinburgh University Press.

Millar, R. M. (2012). *English historical sociolinguistics.* Edinburgh: Edinburgh University Press.

Miller, J. (1993). The Grammar of Scottish English. In J. Milroy and L. Milroy (Eds.) *Real English: The Grammar of English Dialects in the British Isles* (pp. 99-138). London: Longman.

Milroy, J. (1994). The notion of "standard language" and its applicability to the study of Early Modern English pronunciation. In D. Stein and I. Tieken Boon van Ostade (Eds.), *Towards a Standard English 1600-1800* (pp. 19-29). Berlin and New York: Mouton de Gruyter.

Mitchell, H. (1799). *Scotticisms, vulgar anglicisms, and grammatical improprieties corrected.* Glasgow: Falconer & Willison.

Myers-Scotton, C. (1998). *Codes and consequences: choosing linguistic varieties*. New York: Oxford University Press.

Nevalainen, T. & Raumolin-Brunberg, H. (2003). *Historical Sociolinguistics: Language Change in Tudor and Stuart England.* London: Pearson Education.

Paterson, A. (1976). The poor law in nineteenth-century Scotland. In D. Fraser (Ed.). *The new poor law in the nineteenth century* (pp. 171-193). London: Palgrave.

Pietsch, L. (2005). 'Some do and some doesn't': Verbal Concord Variation in the North of the British Isles. In B. Kortmann, T. Herrmann, L. Pietsch and S. Wagner (Eds.). *A Comparative Grammar of British English Dialects. Agreement, Gender, Relative Clauses* (pp. 125-210). Berlin: Mouton de Gruyter.

Sinclair, J. (1782). *Observations on the Scottish Dialect*. London: W. Strahan & T. Cadell.

Sinclair, J. (1831). *The Correspondence of the Right Honourable Sir John Sinclair, Bart: With Reminiscences of the Most Distinguished Characters who Have Appeared in Great Britain, and in Foreign Countries, During the Last Fifty Years. Illustrated by Facsimiles of Two Hundred Autographs..* (Vol. 2). London: H. Colburn & R. Bentley.

Society for the Benefit of the Sons and Daughters of the Clergy. (1845). *The new statistical account of Scotland*. Edinburgh: William Blackwood and Sons.

Watts, R. J. (2012). Language myths. In J. M. Hernández-Campoy and J. C. Conde-Silvestre (Eds.). *The handbook of historical sociolinguistics* (68th ed., pp. 585-606). Hoboken: John Wiley & Sons.

Corpus references

ScotPP = *Scottish Poor Petitions Corpus*. Compiled by Moragh Gordon, Jelena Prokic, Hester Groot, and Alma Strakova [in preparation].

CMSW = *Corpus of Modern Scottish Writing*. Compiled by John Corbett, Jeremy Smith, Wendy Anderson, Jennifer Bann, David Beavan, and Jean Anderson (2007). Available online at https://www.scottishcorpus.ac.uk/cmsw/.

Appendix 1. Proscribed Scotticisms

A compilation of Scotticisms found in the works of David Hume (1752), James Beattie (1787), John Sinclair (1781), and Hugh Mitchell (1799).

Author	Scotticism	Standard English form
Sinclair	a bairn/bearn	a child
Sinclair	a barber	a hairdresser
Sinclair	a bit bread	a bit of bread
Hume	a chimney	a grate
Sinclair	a college	a university
Sinclair	a compliment	a present
Sinclair	a cotter	a cottager
Beattie	a drink	a draught
Beattie	a faint	a fainting fit
Beattie	a flower	a bouquet
Sinclair	a friend	a relation
Sinclair	a good hand of writ	a good hand-writing
Sinclair	a good stipend	a good living, cure, or benefice
Sinclair	a great many company	much company/a great many people
Sinclair	a heritor	a proprietor
Beattie	a hook	a sickle
Beattie	a labouring	a farm
Sinclair	a laird	a squire/lord of a manor
Sinclair	a mains	lands near a mansion-house
Sinclair	a man's subjects	a man's goods, effects
Hume	a park	an inclosure
Hume	a pretty enough girl	a pretty girl enough
Sinclair	a process	a suit, or action at law
Sinclair	a reduction	a suit for reducing
Beattie	a sore head	a headache
Sinclair	a stair	stairs
Sinclair	a tacksman	a leaseholder/ tenant/farmer
Sinclair	a tutor/curator	a guardian
Hume	a wright	a carpenter
Sinclair	a writer	an attorney/ solicitor (should be, author)
Beattie	a yard	a garden
Sinclair	a young man	a bachelor
Beattie	abbacy	abbey
Beattie	Aberdeen's Journal	Aberdeen Journal
Beattie	abort	to miscarry, have an abortion
Sinclair	above	over
Sinclair	abundance	sufficient, enough (in England, plenty/ exuberance)

Beattie	account	bill	Mitchell	ask at someone	ask of
Beattie	accuse one for a crime	accuse one of a crime	Beattie	at six years old	at the age of six
Sinclair	acquaint	acquainted	Sinclair	at worst	at the worst
Beattie	adduce evidence	bring evidence	Sinclair	attour	over and above
Hume	advert to	attend to	Sinclair	aye	always
Mitchell	ailing	sickly, in bad health	Mitchell	badly	sickly, in bad health
Sinclair	albeit	although	Hume	bankier	banker
Sinclair	allenarly	solely, only	Beattie	baxter, brewster, dyster, webster	baker, brewer, dyer, weaver
Beattie	almost never	seldom or never.	Sinclair	be prevailed with to do	be prevailed upon to do
Hume	alongst	along (Yet the English say both amid and amidst, among and amongst.)	Sinclair	be provided in a living	be provided with a living
Sinclair	altogether	in all	Beattie	bear and boar	bear and boar
Hume	alwise	always	Beattie	beast	(not applied to insects, etc.)
Hume	amissing	missing	Sinclair	bedfast	confined to bed/bedrid
Sinclair	an account	a bill	Beattie	behind, i.e. not on time	late, too late
Beattie	an ashet	a plate, at table	Sinclair	below	under
Mitchell	an house	a house	Beattie	below	under
Sinclair	an indweller	an inhabitant/ inmate	Sinclair	ben	in, or into
Sinclair	an old wife	an old woman	Sinclair	better as	better than
Beattie	anent	with regard to: concerning	Mitchell	bid me go there	bade me go thither
Beattie	angry at him	angry with him	Hume	big coat	great coat
Hume	annualrent	interest	Hume	big with a man	great with a man
Beattie	appreciate	appraise	Beattie	black sugar	licorice juice
Hume	as ever I saw	as I ever saw	Sinclair	blyth	gay, merry
Hume	as I shall answer	I protest or declare	Sinclair	body	soul, creature
			Beattie	broke	broken

Beattie	burial	funeral
Beattie	bursar	[student term]
Beattie	burst for laughing	with laughing
Sinclair	but	without
Hume	butter and bread	bread and butter
Beattie	by-table	side-board, side-table
Sinclair	bygone/bypast	past
Beattie	cast up a fault	upbraid one with a fault
Sinclair	catched	caught
Beattie	catholicon	specific
Hume	cause him do it	cause him to do it (yet it is good English to say, make him do it)
Beattie	cautioner, caution	surety, bail
Sinclair	cess	King's land tax
Beattie	challenge, quarrel	reprove, rebuke
Beattie	chamberlain	steward
Beattie	chapman	seller of small wares
Beattie	chimney	grate, iron frame
Beattie	clattering	chattering, chatting
Beattie	clean plate	a plate [at table]
Sinclair	clear (when applied to solids)	bright
Sinclair	clever	quick/active/handy
Beattie	close the door	shut
Sinclair	cloth (in, a cloth-brush)	clothes
Beattie	coarse day	bad day
Beattie	cognosce	take cognizance of
Beattie	come here	come hither
Sinclair	come in by	come in/draw near
Hume	come in to the fire	come near the fire
Hume	common soldiers	private men
Hume	compete	enter into competition
Beattie	complainer	complainant
Sinclair	compleat	complete
Beattie	compliment vs present	compliment vs present
Beattie	conceived in the following words	containing the following words
Sinclair	conform	conformable, according to
Sinclair	connection	connexion
Hume	contented himself to do	contented himself with doing
Beattie	conveyance	to convey
Beattie	corn the horses	feed
Beattie	cousin Germans	cousins German
Beattie	cripple	lame (cannot be an adjective)
Hume	cry him	call him
Sinclair	curt	brief
Hume	cut out his hair	cut off his hair
Hume	debitor	debtor
Beattie	deburse	disburse

Beattie	decreet	decree
Beattie	it was dedicate	dedicated
Hume	deduce	deduct
Beattie	defender	defendant
Hume	defunct	deceased
Beattie	delate	accuse judicially before a court
Sinclair	delicate	sickly, weakly
Beattie	demit an office	resign
Hume	denuded	divested
Mitchell	dependent upon	dependent on
Sinclair	descendents	descendants
Beattie	desire my servant to speak to me	tell him I want to speak to him
Sinclair	desireable	desirable
Beattie	desuetude	disuse
Beattie	detract	take from another's reputation
Beattie	difficulted	puzzled, at a loss
Beattie	disabuse	undeceive
Beattie	discharge	forbid
Sinclair	discreet	civil/obliging
Hume	discretion	civility
Beattie	disposition	writing by which property is transferred
Sinclair	distressed	pained
Beattie	doer	agent
Beattie	draw the table	clear the table
Hume	drunk, run	drank, ran
Hume	dubiety	doubtfulness

Beattie	dull	deaf
Beattie	dure	hard, difficult
Hume	effectuate	effect
Sinclair	eik	also
Sinclair	else	already
Sinclair	eneugh	enough
Sinclair	evenly	even
Sinclair	ever a	any
Hume	evite	avoid
Hume	exeemed	exempted
Beattie	exerce	exercise
Sinclair	expiry	expiration
Beattie	factor	steward
Beattie	failing him	in default of him
Sinclair	familys	families
Sinclair	farm	rent in grain
Beattie	fee	wages
Beattie	feel a sweet smell	smell a sweet smell
Beattie	find no pain	feel no plain
Sinclair	fodder is plenty	fodder is plentiful/ abundant
Beattie	fog	moss
Beattie	follow out a plan	execute, carry on
Beattie	foot of the table	lower end
Beattie	for common	commonly
Hume	for my share	for my part
Hume	for ordinary	usually
Hume	forfaulture	forfeiture
Beattie	foursquare	square

Hume	fresh weather	open weather
Beattie	friend	relation
Hume	friends and acquaintances	friends and acquaintance
Beattie	frost	ice
Beattie	gear	wealth, riches
Beattie	gentlemanny	gentlemanly
Beattie	give me it	give it me
Beattie	go the day	to day
Sinclair	go to the church	go to church
Beattie	go without	unless
Beattie	goat milk	goats milk
Beattie	gown and bands	in a gown and band
Beattie	gravy	fauce
Beattie	great odds	a great change
Beattie	greed	greedy, greediness
Beattie	gutter	dirt
Beattie	half six	half past five
Beattie	half-nothing	less than nothing
Beattie	hard fish	dried fish
Beattie	harvest	autumn
Beattie	has been a strong	must have been
Beattie	hatred at	hatred to, against
Sinclair	he behoved to do it	it behoved him
Beattie	he came again him	against him
Beattie	he dedicate	he dedicates
Beattie	he fevered	he had a fever

Beattie	he is a widow	widower
Sinclair	he will some day repent it	he will one day repent it
Sinclair	he wrote me	he wrote to me/wrote me a letter
Beattie	head of the tabel	upper end
Hume	heritable	hereditary
Beattie	him and me	he and i
Sinclair	hinder to do	hinder from doing
Beattie	hirer	lends a horse for hire
Beattie	hog	swine
Beattie	house to set	to let
Sinclair	how soon	as soon as
Beattie	I am hopeful that	I hope that
Beattie	I asked at him	I asked him; I asked of him
Beattie	I can instruct it	I can prove it
Beattie	I have no fault to him	I find no fault with him
Sinclair	I have no fault to him	I have no fault with him
Sinclair	I love for to do good	I love to do good
Beattie	I stuck among the snow	in the snow
Sinclair	I was not so well last winter	I was not well last winter

Sinclair	I will, he shall	I shall, he will
Sinclair	ilk	each, every
Beattie	implement a	fulfil, perform

	promise	
Hume	in favours of	in favour of
Beattie	in life	alive
Beattie	in my offer	in my choice
Hume	in no event	in no case
Beattie	in place of	in the place of
Sinclair	in place of	instead of
Beattie	in the forty-five	in 1745
Hume	in the long run	at long run
Hume	in time coming	in time to come
Beattie	in use to	used to
Hume	incarcerate	imprison
Sinclair	incomfortable	uncomfortable
Beattie	indeed no	no indeed
Sinclair	independent of	independent on
Beattie	indweller	dweller
Beattie	infeft, infeftment	enfeoff, enfeoffment
Beattie	iniquious	iniquitous, unjust
Beattie	inkholder	inkhorn
Beattie	interlocutor	interlocutory sentence
Sinclair	Is he in?	Is he within?
Beattie	it hurted me	hurt
Sinclair	just so	true, it is so
Sinclair	keept/keeped	kept
Beattie	ken	know
Sinclair	kindling	coals, live coals, firing

Beattie	labour	till the ground
Sinclair	langsyne	long since
Sinclair	large	plentiful, in plenty
Sinclair	last harvest	last autumn
Beattie	lawful day	week day
Beattie	lay your account with opposition	expect opposition
Beattie	lays	lies
Hume	learn	teach
Beattie	libel	indictment
Beattie	liberate	set at liberty
Sinclair	liferent	annuity
Sinclair	light-headed	giddy/delirious
Sinclair	likeways	likewise
Beattie	lime	mortar
Beattie	linens	linen
Beattie	logicks	logick
Hume	lookt over the window	lookt out at the window
Beattie	lost in the river	drowned
Hume	maltreat	abuse
Beattie	marrows	fellows
Hume	marry upon	marry to
Beattie	meat	flesh-meat
Beattie	mercat	market
Beattie	militate against	make against
Beattie	milk-cow	milch cow
Sinclair	misfortunate	unfortunate
Hume	misgive	fail
Beattie	misguides	abuses

Beattie	monday next	Monday nearest
Beattie	mortification	permanent fund
Beattie	mother-in-law	step mother
Beattie	napkin	handkerchief
Beattie	narrate	relate
Beattie	neck	cape (of a coat)
Sinclair	no	not
Beattie	no a good day	not a good day
Sinclair	no more	no farther, only
Beattie	none of them things	those things
Beattie	nonjurant	nonjurer
Sinclair	not so soon	not yet
Beattie	nothing ado	nothing to do
Hume	nothing else	no other thing
Hume	notour	notorious
Sinclair	notwithstanding of that	notwithstanding that
Beattie	oaken deal	plank
Sinclair	oeconomy	economy
Sinclair	oldish	elderly
Hume	on a sudden	of a sudden
Beattie	on the morn	on the morrow
Sinclair	once in the week/year	once a week/year
Beattie	one [...] he, himself	one, oneself

Sinclair	one of these days	one of those days
Sinclair	one would readily imagine	one would naturally imagine
Beattie	onerous	burdensome
Beattie	operate payment	procure or force payment
Sinclair	or then	before then
Sinclair	ordinance	ordnance
Beattie	ornate	elegant
Sinclair	otherways	otherwise
Sinclair	our whole actions	all our actions
Sinclair	out of hand	immediately
Sinclair	overly	carelessly, superficially
Beattie	pannel	prisoner
Hume	paper, pen, and ink	pen, ink, and paper
Beattie	park	inclosure
Hume	part with child	miscarry
Beattie	pen	quill
Beattie	peny, penies	peny, pence
Hume	pepper and vinegar	vinegar and pepper
Beattie	piece cheese	piece of
Beattie	pint	[different amounts]
Sinclair	pitiful	piteous
Beattie	play cards	play at cards
Hume	pled	pleaded
Sinclair	pointed	punctual, accurate

Beattie	policy	pleasure-grounds
Beattie	potage	broth
Beattie	pouch	pocket
Hume	prejudge	hurt
Sinclair	presently	now, at present
Beattie	preses	chairman, president
Beattie	pretty	graceful [dignity implied]
Beattie	process	lawsuit
Beattie	prognostication	almanack
Beattie	propale	divulge
Beattie	prospect	perspective
Beattie	proven	proved
Beattie	pull up by the roots	pluck up
Beattie	purchase	to find, obtain, acquire
Beattie	pursuer	plaintiff, prosecutor
Beattie	queer	comical, humorous
Hume	readily	probably
Hume	rebuted	discouraged by repulses
Beattie	reckon	think
Sinclair	reconciled with	reconciled to
Beattie	relevant	sufficient, valid
Beattie	remeed	remedy

Beattie	repeat a sum paid	repay
Beattie	restrict	limit, confine
Beattie	rests me nothing	owes me nothing
Beattie	roasted	toasted
Sinclair	run, drunk	ran, drank
Beattie	say the grace	say grace
Beattie	scarce of	short of
Sinclair	Scots	Scottish, Scotch
Beattie	seed	saw
Beattie	seeking his meat	begging
Hume	severals	several
Beattie	shearers	reapers
Beattie	sheriff depute	deputy sheriff
Beattie	shirt	shift
Sinclair	sib	akin
Sinclair	sicker	sure, certain
Beattie	simply	absolutely
Sinclair	sirname	surname
Sinclair	situate	situated
Sinclair	slidderly, slippy	slippery
Beattie	slipped a foot	my foot slipped
Mitchell	so as	so that
Sinclair	so soon as	as soon as
Hume	some better	something better
Beattie	some day	one day
Beattie	sparse	loose
Beattie	speak to me	speak
Beattie	spice	pepper

Sinclair	split new	quite new
Beattie	stair	stairs
Beattie	stammers	stumbles
Beattie	stay	lodge, live, dwell etc.
Beattie	stingy	penurious
Beattie	stipend	salary
Beattie	storm	[any elements]
Beattie	subjects	effects
Beattie	subsists	supports, maintains
Beattie	the sugar, the rum	[no article]
Beattie	sunday's morning	sunday morning
Hume	superplus	surplus
Beattie	sustain	admit
Beattie	sweet	fresh
Beattie	tags of a shoe	straps of a shoe
Beattie	take on	inlift
Beattie	tea-kitchen	tea-urn
Sinclair	teach't	taught
Hume	tear to pieces	tear in pieces
Beattie	tell him to	bid him come
Beattie	tempt'd	tempted
Beattie	tender	sickly, weakly
Hume	tenible argument	good argument
Sinclair	the [year]	[year]
Beattie	the better of	the better for
Beattie	the botany	botany
Beattie	the cold	a cold
Sinclair	the more, that	the more, as/because
Sinclair	the ordinance	the sacrament/eucharist
Beattie	the piece	a piece, each
Beattie	the shore	quay, wharf
Beattie	there is no matter	no matter, it is no
Hume	there, where	thither, whither
Beattie	thereafter	after
Beattie	thereby	thereabout(s)
Beattie	thinks long	longs for
Beattie	this here man	this man
Sinclair	this much, and that much	thus much, and so much
Beattie	throng	full
Sinclair	through	across
Beattie	through his sleep	in his sleep
Hume	Thucydide, Herodote, Sueton	Thucydides, Herodotus, Suetonius
Beattie	timber	wooden
Beattie	timeous	temely
Sinclair	timously	timely, early
Hume	tis a question if	tis a question whether
Hume	tis a week since he left this	tis a week since he left this place
Beattie	to a wish	according to our wishes

Sinclair	to affront any one	to get the better of any one
Sinclair	to ask/inquire/demand at a man	of a man
Sinclair	to be a missing	to be missing
Hume	to be angry at a man	to be angry with a man
Hume	to be difficulted	to be puzzled
Beattie	to be found	to [letter address]
Sinclair	to be liable in a compensation	to a compensation
Sinclair	to call for a person	to call on a perosn
Sinclair	to cause a person	to make a person
Beattie	to condescend upon	to specify
Hume	to crave	to dun, to ask payment
Sinclair	to demit	to resign
Hume	to depone	to depose
Hume	to discharge	to forbid
Sinclair	to do any thing to purpose	to the purpose / to good purpose
Sinclair	to do bidding	to do what is bidden
Hume	to enquire at a man	to enquire of a man
Beattie	to evite	to shun, avoid
Beattie	to exeem	to exempt
Hume	to extinguish an obligation	to cancel an obligation
Sinclair	to fever	to be seized with a fever
Hume	to furnish goods to him	to furnish him with goods
Hume	to get a stomach	to get an appetite
Sinclair	to give one a hat	to make a bow to any one
Sinclair	to go out walking	to go out a walking
Beattie	to homologate	to ratify
Sinclair	to learn	to teach
Sinclair	to lodge in a house	s
Sinclair	to make of one	to make much of one
Sinclair	to make songs of one	to praise one much
Beattie	to mandate	commit to memory
Sinclair	to mind	to remember
Beattie	to my bed	to bed
Beattie	to notice	to take notice of, to mention
Hume	to open up	to open, or lay open
Beattie	to plenish	to furnish
Hume	to remeed	to remedy
Sinclair	to restrict	to limit/confine
Sinclair	to roar	to cry/weep

Sinclair	to rove (in a fever)	to be light-headed/delirious
Sinclair	to see about one	to see/inquire/look after one
Hume	to send an errand	to send of an errand
Sinclair	to set a house	to let a house
Sinclair	to spier	to ask/enquire
Sinclair	to stay in a house	to reside in a house
Sinclair	to steik a door	to shut a door
Beattie	to succumb	to sink under
Beattie	to summons	to summon
Hume	to take off a new coat	to make up a new suit
Sinclair	to tell upon one	to tell of one
Sinclair	to think shame	to be ashamed
Sinclair	to wait on a person	to wait for a person
Sinclair	to want for any thing	to be without any thing not desirable
Sinclair	tofore	before
Beattie	took the pox	was seized with the small pox
Sinclair	toply	finely
Beattie	topped	tapped
Beattie	tradesman	shopkeeper
Beattie	transported	translated
Beattie	turned sick	became sick
Sinclair	twenty years or	or thereabout

	thereby	
Sinclair	two weeks	a fortnight
Beattie	unformal	irregular
Beattie	up streets	up the street
Hume	vacance	vacation
Beattie	versant	conversant
Beattie	victual	corn
Beattie	vivres	provisions
Beattie	vocable	word
Beattie	wait of you	wait on you
Beattie	was made do it	was made to do it
Beattie	watch is behind/before	slow/fast
Beattie	water of dee	river Dee
Beattie	weaving, working	knitting
Sinclair	wee	little
Beattie	what airt is the wind in	how is the wind?
Sinclair	whenever	as soon as
Beattie	wife [as, old woman]	wife [with man]
Beattie	will	shall [p. 27-28]
Beattie	will not readily	is not likely to
Hume	with child to a man	with child by a man
Beattie	witnessed	beheld, saw
Beattie	would die before I would break	rather than break
Beattie	would have you to know	have you know

Beattie	wrongous	wrongful
Beattie	wrote him	wrote to him
Hume	yesternight	last night
Sinclair	yon, yond	that
Beattie	you eat little yesterday and have eat nothing to day	ate little, have eaten [preterite, passive participle]
Sinclair	you was	you were
Sinclair	your favours	your favour
Sinclair	youthy	youthful

Appendix 2. Scotticisms and Standard English forms in ScotPP corpus

Scotticism	Standard form	Scotticism occurrences	Standard English occurrences
[singular]	for [plural]	That the petitioner haves presented; That your petitioner, being confined & afflicted for three years by a painful disease, and one that has rendered me now quite helpless, find it; That the Petitioner is in indigent circumstances have five of a weak family	
a bairn/bearn	a child	Likewise I wish to say that my wife would rather take to two Bairns	in behalf of the Child; The Board being pleased at that time to suspend taking the Child upon the Roll, untill such time as an Examination of the parants of the Child would take place; Grant the Child to be put upon the Roll; for Sinclair's Child expects very soon to be delivered of a child
a cotter	a cottager	The Petition of John McKay alias Donn Cotter	
a friend	a relation		Her relations are neither able nor willing to assist her
a process	a suit, or action at law	as the Expense of this Procefs; I hope you will grant me the Procefs	the Petitioners are unable to bear the expence of a Law Suit; an action being raised; in the present action; the necessity of instituting an action against him; instituting an action at her

			instance; about to raise an action at her instance
a writer	an attorney/solicitor (should be, author)	at the instance of John Gowans Writer a settlement with Mr Gowans writer	
account	bill	said to be an account; the Receipt for the same the other parts of the account	Settlement with the two Bills; the amount of the Bills
acquaint	acquainted		she believes that your Board is already so well acquainted with her case
alwise	always		but would require always a person to take care of her
anent	with regard to		1st with Regard to the account the defender has only knowledge of two parts of it
be prevailed with to do	be prevailed upon to do	her Mother who cannot be prevaild with to remain	the Pursuer was prevailed upon
bedfast	confined to bed/bedrid	I am most of my time Bedfast; has been for the most part bedfast; has him bedfast for the last 9 months; That the petitioner is Bedfast for a considerable time since	That the Petitioner Aged 86 years has been bedridden; and is now almost confined to bed.; I have been altogether confined to Bed; and is now Bed ridden and greatly sufferring; she has been bedridden.; That the Petitioner is Bedridden since 6 weeks back
decreet	decree	Extracting the Decreet to follow herew.; ordained by Decreet of Court	
defender	defendant	although he has got in defences from the defender; 1st with Regard to the account	

		the defender has only knowledge of two parts; Therefore the said Robert..Kerr Defender	
desire my servant to speak to me	tell him i want to speak to him	yet he refuses altho' often desired & required to do so	
distressed	pained	in my present Situation in poverty and distres of body	
do not mind	do not attend to		I require now a Person to attend me when out of bed; My mother alone has to attend to me; and requires always me to attend on; her only daughter Betty who cannot leave the house to earn any thing for herself but must attend on the Petitioner him
dull	deaf		who is both blind, deaf and dumb
fee	wages	additional aid to pay Doctors fees	
follow out a plan	execute, carry on		Being deprived of executing summons for the Board
he dedicate [singular]	he dedicates [plural]	Therefore your humble Petitioner Expect	
he is presently in London	he is now/at present in London		therefore the Petitioner is at present going from house to house seeking shelter
i am hopeful that	i hope that		Therefore I hope that the Board will take my our case; I hope that the Honourable Members will be pleased to consider your petitioner.s. request; she hopes that adequate

			support will be granted; I hope that the Members of the Board will be pleased to grant; I do not expect to trouble you long & hoping you may be pleased to give my case a favourable consideration, your petitioner; which she hopes the Honourable Board will be pleased to grant her; I hope that the Honourable members will be pleased to consider; she hopes that that your honourable board will grant a more liberal sum; therfore she hopes your Honourables will take her "Case" into Consider--ation
i can instruct it	i can prove it	which is instructed by the certificate produced	
I will, he shall	I shall, he will	the petitioner shall every pray; if you don't comply to my necessities I will appeal to some other Churt; he shall reckon it a favour conferred upon him	the Cot being also too small and will let in the rain; the Honourable will be pleased to allow a little addition; I hope that the Honourable Members will be pleased to consider; [he] has stopt short &will pay no more; as your Lordship will see
keept/keeped	kept	also that one shilling has been keept off this month; but has stopt short &will pay no more	
lays	lies		which prevents her from lying down at night

libel	indictment	also stated in her Lible; Refuseing to obtain the Libel	
maltreat	abuse		That the Petitioners were lately assaulted abused and "baghashed"
our whole actions	all our actions	being deprived of the whole of her Property	
presently	now, at present	presently Private in the Royal Perthshire Mekka is the father; presently residing at Logierait	the object that this petition is given in for, is at present -35- years of Age; at Present she stands very much in need of a Gown; at Present it may serve for a year; I Live with my Brother at present
process	lawsuit		And as she is unable to bear the expence of a Lawsuit
proven	proved		which can be proved by investigation
pursuer	plaintiff, prosecutor	to the Pursuer of £2 Sterling. of inlying expenses,	
reckon	think		
restrict	limit, confine		the said office being confined on a man from another parish
stay	lodge, live, dwell etc.		That thi Hut in which the petitioner dwells is unfit for any human being to live in
subsists	supports, maintains	has neither stock nor other means of subsistance	cannot support herself; and the land was making some support for me; unable to support herself by labour
the [year]	[year]		it does not meet the one half
this much, and that	thus much, and so much		who is evidently so much reduced by Rheumatic attacks

much			
to crave	to dun, to ask payment	that she has no relation from whom she can crave shelter	
to my bed	to bed	and was in consequence confined to her Bed	
to stay in a house	to reside in a house		John Constable residing in Stanley; presently residing in the South Street; presently residing at Logierait
to want for any thing	to be without any thing not desirable	The Petitioner would most humbly and earnestly solicit the attention of the honourable Parochial Board to her present want; I am sory to say that it does not meet the one half of our pecuniary wants.	
to want it	to be without a thing, even though it be not desireable		quite destitude having no means of support & without any to assist them

Appendix 3. Scotticisms and Standard English forms in CMSW dataset

Scotticism	Standard form	Scotticism occurrences	Standard English occurrences
a friend	a relation		a comparision which would be desirable on a question among near relations
a labouring	a farm		I have got a very extensive farm from my noble master; when laying in stock for my farm
a writer	an attorney/solicitor (should be, author)		Genl. Balfour speaks in one of his letters of a Mr. [¿] as the Attorney employed
an account	a bill	he had charged me in his account	For God's sake Mr Murray accept the bill; at bills of six nine or twelve months
anent	with regard to		With regard to my reception of Mr John Hurst; particulars with regard to my health; I dont know how to do in regard to my servants; & thank you for the attention you have paid to my donations in regard to the money
anent	concerning		concerning which I have, for the present, only to repeat my urgent request
as ever I saw	as I ever saw	time flies faster in this capital than ever I have found it	
body	soul, creature	I would not mention my new work to any other body	
bygone/bypast	past	- The one in bygone London the other I can say little of	

desire my servant to speak to me	tell him i want to speak to him	desired Mr Wardlaw to send a copy	just tell her what she must do
doer	agent		I have followed your example in chusing an agent
effectuate	effect (This word in English means to effect with pains and difficulty)		I am printing in the most beautiful manner that modern art can effect
gentlemanny	gentlemanly		the most amiable and gentlemanly Robt Crawford Esqr; your most gentlemanly friend Lord Clair; He is a very Gentlemanly young man
have been badly; poorly	ill, fickly, in bad health		Constable is very ill.
he is presently in London	he is now/at present in London		I have not time to enlarge more at present; little William is well and sleeping at present; I have very much need of it at present.; which at present fall upon me.
i am hopeful that	i hope that		I hope to see you on Sunday; I hope you are well; I hope you got safely to Sheraton and have been well since; I hope you will tell me when you see it; I hope the paper in Frazer will do no harm
I behoved to	it behoved me		if you think it behoves you to do

go	to go		so
i can instruct it	i can prove it	Mr Hobhouse is sorry that this instruction is contrary to the spirit of the Act	
I will, he shall	I shall, he will	let it be in a kind letter else I will be exceedingly grieved; I will cause O. & B. to account for it at your next settlement	Our mutual friend Sir Walter Scott will have been telling you
ilk	each, every		one inch & a half on each side of the red line
independent of	independent on	(independent of the pecuniary advantages)	
keept	kept	my [?] as a judge will fall to be stopt	
lays	lies		The country beyond is elevated and lies in three terraces; It lies or rather hides itself in the water
likeways	likewise		I have followed your example in chusing an agent and [gone] likewise to Oliver & Boyde.; and perhaps likewise that I would be straitened of money; The Edition of the Songs with notes consists likewise of 1000
mind it	remember it		You may remember that Cairoli in a measure saved the Kings life
propale	divulge		He advised me not to do it nor to divulge the trifling matter to any one
purchase	to find, obtain, acquire		he never acquired any other domicil
rests me nothing	owes me nothing		whether he had owed me ought or not if

			the lion encountre is absolutely abominable
simply	absolutely		
stay	lodge, live, dwell etc.	so that you can come and stay with me	Two Dutch clergymen at the Cape were once lighted on comfortably lodged
stipend	salary	— he makes an assertion that the stipend is so much	
sustain	admit		Admitting that while he remained in India in the Kings Service Scotland
the piece	a piece, each		give Captain Need & Dr Smith a copy each for their kindness
think shame	be ashamed	I think shame for having been so long in writing to you	
to call for a person	to call on a perosn		I called on Mr. Law but he had no copy
to condescend upon	to specify		you need not specify what she has not to do,
to lodge in a house	to live		lived in it the rest of the year
to notice	to take notice of, to mention		he takes no notice of it in his opinion; I would not mention my new work to any other body.
to spier	to ask/enquire		When you asked about appendices; you may ask Vardon if you doubt; I would ask you to leave out my observation book; and I have asked her to present the other to the Queen of Italy.; The conditions between two such men should never be enquired about;

to tell upon one	to tell of one		who had to tell her of it
to want it	{ to be without a thing, { even though it be{ not desireable	yet wanting I Armenian Epistles; for there is a deplorable want of books at Rome; It only wants binding.	
two weeks	a fortnight		for a fortnight at one time; sent by steam I shall have it in a fortnight; I was a fortnight in Rome on my way south to Naples; and intend to stay another fortnight.
was married on	married to		was married a few days ago to Miss Parnell
will not readily	is not likely to		when you see it whether you think it likely; But in general we prefer Mr Moncrief as being likely to go more to the bottom of such a question; & I believe it will appear that they say that I was not likely to get the better of this

Appendix 4. Non-proscribed Scotticisms in the ScotPP and CMSW

Non-proscribed Scotticisms occurring in the ScotPP dataset. **Scotticism**	**Standard form**	**Scots occurrences**	
close	finish	1	● till it come to aclose (Margaret Spence, undated [1821])
has not	hasn't	40; or, all instances save 2 (i.e. *n't* occurred just twice)	
never (as past negation)	not	3	● he never at any time said to her that such a sum was due (Margarete Spence, 13 January 1821) ● she never troubled them while she possessed anything herself (Scribe for Sophia McKay, 3 February 1852) ● your Petitirner, who never had land, is for more than 20 years… (William McKay, undated [1851])
should	ought	3	● [we] should be thankful for the little that the Board is giving (John Donn, 14 December 1851) ● her case should not be neglected (R.W. Black, 16 March 1852) ● it should have been increased (George Mackay, 13 August 1852)
can	may	1	● more than I can expect from those who… (Anne McKay, 31 July 1852)
auld	old	3	● I am an auld frail woman (Isabella McKay, 14 January 1852) ● I am an auld frail woman (Miney Macleod, 3 February 1852) ● Exept another old Frail auld woman (Elssy McKay, 28 December 1851)

Non-proscribed Scotticisms occurring in the CMSW dataset.

Scotticism	Standard form	Scots occurrences	
did not	didn't	94; or, all instances	
never (as past negation)	not	1	● Never a blessed word have we got from any of our friends since we left England (David Livingstone, 1 November 1859)
undertaken		4	● though I could not undertake it myself , you might find some writer to undertake (Samuel Smiles, 26 November 1859) ● & Mrs Greig has kindly undertaken to correct for me (Mary Somerville, 3 April 1842) ● Admiral Smythe's son in law has very promptly undertaken the editorship (David Livingstone, 24 August 1857)
forenoon	late morning	1	● come with the forenoon omnibus if in time (Mary Mather, April 29 1841)
advocate	barrister	1	● I have since seen the Advocate (Archibald Campbell, 2 February 1825) ● the Advocate [Onesulent] & also D Melville having held Action (Archibald Campbell, 2 February 1825)
should	ought	6	● The conditions between two such men should never be enquired about (James Hogg, 24 March 1821) ● Black-wood says a fair medium price should be named and agreed on (James Hogg, 5 May 1821) ● I do not think Miſs Balfour should administer till the case is decided (Archibald Campbell, 4 May 1823) ● no time should be lost in getting the Warrant (Archibald Campbell, 22 December 1823) ● which as a general rule I may mention should extend one inch & a half… (Burnes, 28 December 1834) ● I have thought an antelope named Nakon should be introduced into one of Wolf's pictures (David Livingstone, 5 March 1857)
want binding		1	● It only wants binding. (David Livingstone, 5 March 1857)

needs no counting		1	● The last edition of the Wake needs no counting (James Hogg, 24 March 1821)