



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

A Study of *Gan*, *Can* and *Beginnen* in the Northern English and Scots of the Late Fourteenth and the Fifteenth Centuries

Wojciech Gardela

Doctor of Philosophy
The University of Edinburgh
2017

I declare that this study:

- has been composed by me;
- is my own work; and
- has not been submitted for any other degree or professional qualification.

Abstract

In Middle English and Scots, instances of **gan** and **can** behave differently from etymologically related **beginnen** in that they are mainly, or exclusively, found with the plain infinitive and with a non-ingressive meaning. They also occur in narrative verse (rhymed and non-alliterative), where they have a metrical, intensive-descriptive or textual function. All of this suggests that **gan** and **can** are more advanced in the divergence of their development towards auxiliarity than the verb **beginnen**. Earlier studies mainly concentrate on the meaning and/or function of **gan** and **can** in verse (Wuth 1915, Beschorner 1920, Funke 1922, Mustanoja 1960, Kerkhof 1966, Visser 1969 and Brinton 1981; 1983; 1988 amongst others), whereas investigations by Brinton (1981; 1988; 1996), Ogura (1997; 1998; 2013) and Sims (2008; 2014) address the divergence in the development of this verb and its variant in terms of grammaticalization, but with references to Middle English in general. Studies by Los (2000; 2005), on the other hand, deal with the grammaticalization of **onginnan** and **beginnan** with the plain infinitive in Ælfric's works. However, no studies have been carried out on whether **gan** and **can**, as well as **beginnen** develop differently in terms of grammaticalization in the 'English' of the six northern counties of England and of Scotland in the late 14th and the 15th centuries, conventionally referred to as Northern Middle English and Early Scots, respectively. With the aid of Northern Middle English and Early Scots texts from computerised corpora (*The Helsinki Corpus of English Texts*, *The Innsbruck Corpus of Middle English Prose* and *The Teaching Association for Medieval Studies*, as well as *The Helsinki Corpus of Older Scots* and *A Linguistic Atlas of Older Scots*), this study looks into whether: a) **gan** and **can**, as well as **beginnen** differ with respect to their morphological paradigms, in view of what we know about grammaticalization and the development of invariant forms? b) these verbs differ with respect to their complements, in view of claims in the literature that the more grammaticalized variant takes the plain infinitive; and c) **gan** and **can** are a development from **onginnan** and **aginnan**, originally expressing ingression but shown in the literature to have undergone semantic bleaching in Old English and in early Middle English period? This study shows that in Northern Middle English and Early Scots,

gan and *can* display characteristics of grammaticalization, while *beginnen* participates in global language changes affecting the category of the verb in ME and Scots.

Acknowledgements

I would like to thank the Arts and Humanities Research Council for their financial support between 2007 and 2014; without it, this study would not have seen its daylight. I would also like to thank the following: Prof. Heinz Giegerich and Dr Chris Robinson for their help with the metrical scansion of lines containing *gan* and *can* in the data sets under investigation; Dr Rhona Alcorn and Dr Meg Laing for their help in translating some Northern Middle English data; for their patience, support and helpful guidance, I am especially grateful to my supervisors: Dr Linda van Bergen and Dr Graeme Trousdale. Thank you also to my Mother, Brian, Irène and Ewan for their continuous support and understanding, and patience, throughout the whole time.

Table of Contents

Abstract	3
Acknowledgements	5
Table of Contents.....	6
Abbreviations	10
Chapter One – Foundations	12
1.1 Introduction.....	12
1.2 Aims and Objectives.....	13
1.3 Northern Middle English and Early Scots as Related Varieties.....	27
1.3.1 Linguistic Divergence within the Northumbrian <i>Sprachgebiet</i>	27
1.3.2 Availability of Written Evidence	31
1.4 Approaches to the Study of <i>Gan</i> and <i>Can</i>	34
1.4.1 Semantic and Pragmatic.....	34
1.4.2 Syntactic	36
1.4.3 Grammaticalization	37
1.5 Scope of the Work.....	39
1.5.1 <i>Gan</i> and <i>Can</i> as Separate Lexemes or Variants of <i>Ginnen</i>	39
1.5.2 Periodisation of the Data Sets.....	43
1.6 Structure of the Thesis	45
Chapter Two – Methodological Issues	47
2.1 Introduction.....	47
2.2 Materials	47
2.2.1 Northern Middle English Data.....	48
2.2.2 Early Scots Data	54
2.2.3 Representativeness.....	57
2.2.4 Spelling Variation	64
2.3 Classification of Items under Investigation	65
2.3.1 Semantic Criteria.....	66
2.3.2 Syntactic Criteria	78
2.4 Research Approaches Used Towards Identifying Items under Investigation ...	81
2.4.1 Quantitative Method	81

2.4.2	Qualitative Method	104
2.5	Summary	106
Chapter Three – Changes in Form		107
3.1	Introduction.....	107
3.2	Changes in the Morphology of Verbs in the Transition from Old English to Middle English	109
3.3	Grammatical Forms and Inflectional Morphology of <i>Gan</i> , <i>Can</i> , and <i>Beginnen</i> in Middle English and Scots	117
3.3.1	<i>Gan</i> and <i>Can</i>	117
3.3.2	<i>Beginnen</i>	129
3.4	Morphological Processes Involved in the Grammaticalization of <i>Gan</i> , <i>Can</i> and <i>Beginnen</i> in Northern Middle English and Early Scots	154
3.4.1	Morphological Behaviour of <i>Gan</i> and <i>Can</i> in Middle English Indicative of Grammaticalization	155
3.4.2	Decategorialization	158
3.4.3	Analogy.....	164
3.4.4	Phonological Reduction.....	167
3.5	Summary	176
Chapter Four – Changes in Complementation		178
4.1	Introduction.....	178
4.2	Changes in the Infinitive Found with Verbs in the Transition from Old English to Middle English	179
4.3	Types of Complements Found with <i>Gan</i> , <i>Can</i> and <i>Beginnen</i> in Middle English and Scots	181
4.3.1	<i>Gan</i> and <i>Can</i>	181
4.3.2	<i>Beginnen</i>	190
4.4	Accounting for Infinitival Complements of <i>Gan</i> , <i>Can</i> and <i>Beginnen</i> in Middle English and Scots.....	198
4.4.1	Metrical and Stylistic Considerations.....	199
4.4.2	Reanalysis and Analogy-Based Syntactic Approach	203
4.5	Summary	231

Chapter Five – Changes in Meaning	233
5.1 Introduction.....	233
5.2 Meaning of <i>Gan</i> , <i>Can</i> and <i>Beginnen</i> in Middle English and Scots.....	234
5.2.1 <i>Gan</i> and <i>Can</i>	234
5.2.2 <i>Beginnen</i>	244
5.3 Determining the Meaning of <i>Gan</i> , <i>Can</i> and <i>Beginnen</i>	249
5.4 Changes in the Meaning of <i>Gan</i> , <i>Can</i> and <i>Beginnen</i> in Northern Middle English and Early Scots.....	261
5.4.1 Semantic Bleaching	261
5.4.2 Opacity	272
5.4.3 Persistence	278
5.5 Summary	280
Chapter Six – Changes in Context	282
6.1 Introduction.....	282
6.2 Context of <i>Gan</i> and <i>Can</i> in Middle English and Scots	283
6.2.1 Text and Linguistic Environment.....	284
6.2.2 Metre and Rhythm	295
6.2.3 Alliteration and Rhyme	304
6.3 Proposals on Function(s) of <i>Gan</i> and <i>Can</i> in Middle English and Scots Verse Texts.....	319
6.3.1 Metrical and Rhyming Function	319
6.3.2 Intensive-Descriptive or Stylistic Function	329
6.3.3 Textual Function.....	331
6.4 Summary	346
Chapter Seven – Conclusions	348
7.1 Meeting Aims and Objectives	348
7.2 Regional Differences in the Distribution of <i>Gan</i> , <i>Can</i> and <i>Beginnen</i> in Northern Middle English and Early Scots.....	353
7.3 Limitations of the Study	354
Appendix One.....	356
Appendix Two	361
Appendix Three.....	362

Bibliography..... 363

Abbreviations

Dialectal and Language Labels:

OE – Old English

eME – early Middle English

IME – late Middle English

ME – Middle English

NME – Northern Middle English

SME – Southern Middle English

EModE – Early Modern English

ModE – Modern English

ESc – Early Scots

MSc – Middle Scots

OSc – Older Scots

Reference Works:

DOE – Dictionary of Old English: A to G Online

DOST – A Dictionary of the Older Scottish Tongue

DSL – Dictionary of the Scots Language

eLALME – An Electronic Version of A Linguistic Atlas of Late Medieval English

LALME – A Linguistic Atlas of Late Mediaeval English

MED – Middle English Dictionary

OED – Oxford English Dictionary

Lexical and Grammatical Categories:

adv. – adverb

comp. – complement

ger. – gerund

Inf – infinitive

n. – noun

p.t. – past tense

v. – verb

AdvP – Adverb Phrase

CP – Complementiser Phrase

IP – Inflection Phrase

NP – Noun Phrase

PP – Prepositional Phrase

V/VP – Verb/Verb Phrase

Other Abbreviations:

cf./Cf. – *confer*

fol. – folio

l. – line/lines

LP – Linguistic Profile (*eLALME*)

MS/MSS – manuscript(s)

n.d. – no date

p./pp. – page/pages

P – part

Pl – plural

Sg – singular

st. – stanza

St. – Saint

s.v. – *sub verbo*

s.vv. – *sub verbis*

vol./Vol. – volume

Chapter One – Foundations

1.1 Introduction

This study deals with the verb *gan* and its variant *can*¹, as well as *beginnen* in the ‘English’ of the six northern counties of England and of Scotland in the late 14th and the 15th centuries, conventionally termed ‘Northern Middle English’ (NME) and ‘Early Scots’ (ESc), respectively.² In this study, I investigate the grammaticalization of *gan* and *can* in the NME and ESc data sets from a historical perspective, i.e. I consider whether Old English (OE) *onginnan* and OE and early Middle English (eME) *aginnan* are the source of the above verb and its variant and I look at typical mechanisms and processes of language change that *gan* and *can* undergo: decategorialization with associated morphological fixation, phonological reduction, analogy, reanalysis and semantic bleaching. In this thesis, I also look into the grammaticalization of *gan* and *can* in the NME and ESc data sets from the pragmatic point of view, where I examine the function(s) of this verb and its variant, as well as the verb in infinitive, in narrative verse. By combining both perspectives on the grammaticalization of *gan* and *can* in the NME and ESc data sets, I aim to show a journey on which this verb and its variant go, namely one towards auxiliation. However, this change that *gan* and *can* undergo from lexical to grammatical would not be complete without references to the etymologically related *beginnen* – I demonstrate in this study that as a marker of ingressive aspect, this verb behaves differently from *gan* and *can*. For example, I show that *beginnen* occurs in finite and

¹ From this point onward, I clarify some terminology in relation to *gan* and *can* for ease of reference. When I use ‘*gan*’ in this thesis, I refer to the third person singular preterite form *gan* with infinitive, unless stated otherwise. Similarly, when I refer to ‘*can*’, I refer to the third person singular form *can* with infinitive, unless stated otherwise. In terms of form and complementation, these represent the auxiliary uses of the verb *gan* and its variant *can* in Middle English and Scots. In addition, although *can* appears to be in the present indicative in form because of its similarity to the modal (auxiliary) verb *can*, it has preterite meaning because it is a variant of *gan* (cf. Section 1.5.1). In this respect, this variant is also treated here as being in the preterite in form.

² In this study, I use labels denoting varieties in two ways: a) to refer to the forms of ‘English’ spoken and written in England or Scotland during the given period; and b) to signify the period when such forms were used. Although I follow the conventionally adopted periodisation of English and Scots given in Macafee (n.d.: §1.1.3), cf. Section 1.5.2 on further information on the dating behind the labels ‘NME’ and ‘ESc’. Since a number of sources referred to in this study use the label ‘Scots’ without providing further particulars, I use this label in a loose way to refer to a variety of English used in Scotland during the time when *gan* and *can* were in use.

non-finite constructions, participating in the global changes related to verbs at the time. I also investigate the complementation of this verb, given the increased use of the *to*-infinitive with non-auxiliary (modal) verbs, instead of the plain infinitive.

This study begins with an explanation of why the topic of whether these verbs develop differently in these varieties is a worthy topic of detailed research, immediately followed by the aim and objectives of this study (Section 1.2). In Section 1.3, I offer a rationale as to why NME and ESc are interesting with reference to the topic of this thesis, but I also highlight the problems with the availability of textual evidence of a specific kind in the predecessors of NME and ESc, namely the Northumbrian dialect of OE and Pre-literary Scots, respectively.³ In Section 1.4, I introduce approaches (semantic, functional and syntactic, as well as from the grammaticalization perspective) adopted by previous studies into *gan* and *can*, as well as *beginnen*, and identify the gap that this study aims to fill. The subsequent section specifies the scope of the work undertaken in this work, more specifically: a) it looks into whether the verb *gan* and its variant *can* should be considered as lexemes or as variants of *ginnen*, and what is the rationale for the treatment of *can* as a variant of *gan* in this study; and b) the reasoning for altering the conventionally accepted periodisation associated with the ‘NME’ and ‘ESc’ labels, and the considerations for dividing the data sets into two periods: 1350/1375-1420, referred to in this study as Period I, and 1420-1500, represented here as Period II. The chapter concludes with a summary of the organisation of the remainder of this study (Section 1.6).

1.2 Aims and Objectives

Although etymologically related to *beginnen*, the verb *gan* and its variant *can* diverge in their development along the grammaticalization cline towards auxiliiation.

³ It is generally acknowledged that 1375 marks the beginning of the literary Scots period (Macafee n.d.: §1.1.3, Macafee 1992/93: §1.4) with Barbour’s *The Bruce*. However, it is possible to find 1380 in some sources, including the *LAOS*, as the onset of this period. This is because Barbour completed the main section of the above work in 1375, with an additional section finished in the 1380s. This work survives in two 15th century MSS by John Ramsay: Edinburgh, National Library of Scotland Advocates 19.2.2 from 1489 and Cambridge, St. John’s College G. 23 from 1487 (imperfect copy starting in Book 4) (J. Smith 2013: 39). In this study, 1375 is treated as the date marking the beginning of the literary Scots period, unless stated otherwise.

There are a number of different interpretations of what the term ‘grammaticalization’ stands for, ranging from the ‘classical’ ones offered by Meillet ([1912] 1958) and Kuryłowicz (1965) to others, with this term being extended in various ways, for example, by Hopper (1991), Matisoff (1991), Traugott & Heine (1991), Lehmann (1995) and Fischer & Rosenbach (2000). In this study, I adopt the view on grammaticalization provided by Hopper & Traugott ([1993] 2003: 1) where this process is seen to be concerned with changes of how lexical items and constructions develop grammatical functions in certain contexts, or how already grammaticalized constructions develop new grammatical functions. In particular, the process of grammaticalization is characterised by a number of cross-linguistically recurring relationships between morphosyntactic, phonological, semantic and pragmatic changes, and so, it is seen as a subset of phenomena occurring in change. It is also gradual, consisting of a number of ‘small’ steps that lead to gradience between different categories (cf. Brinton & Traugott 2005: 26-27, 150, Traugott & Trousdale 2010, Traugott 2012a; 2012b). Hopper & Traugott ([1993] 2003: 2) argue,

[grammaticalization] provides the conceptual context for a principled account of the relative indeterminacy in language and of the basic non-discreteness of categories. As a term referring to actual phenomena of language, ‘grammaticalization’ refers most especially to the steps whereby particular items become more grammatical through time. Grammaticalization in this sense is part of the wider linguistic phenomenon of structuration, through which combinations of forms may in time come to be fixed in certain functions.

Functional variation of *gan* and *can*, as well as *beginnen* is the result of different grammaticalization paths, one of which is auxiliation. Auxiliation is the process involving a morphosyntactic change from the lexical structure of the ‘verb-complement’ type to a grammatical structure of the ‘grammatical marker-main verb’ type. It also involves semantic and phonological changes (Kuteva 2001: 1-2). The success of this change can be seen in the development of an auxiliary-verb system in Middle English (ME) and Early Modern English (EModE), where both lexis and grammar interacted in interesting and complex ways, resulting in a number of auxiliary constructions. According to Warner (1993: 103), verbs characterised by auxiliation (modal and auxiliary verbs) displayed a number of different properties in ME, setting them apart from other, lexical, verbs: a) “[o]ccurrence in ellipsis in the

same way to the modern post-auxiliary ellipsis and pseudogapping”; b) “[o]ccurrence within impersonal constructions, where the subordinate verb controls the case of nominal arguments”; c) “[r]estriction of some of these words to finite forms”; d) “[u]se of past-tense forms without past time reference, outside a motivating context”; e) “[s]ubcategorization for the plain infinitive, not for the to-infinitive”; and f) “[p]reterite-present morphology”⁴. (Warner 1993: 103 also comments that this list shows properties which occur with at least some of the verbs in question).⁵ In particular, therefore, the use of the plain infinitive is typical of modal verbs within an auxiliary group, although there are some exceptions, too. I will come back to this point when I set out the criteria in Section 2.3 for isolating instances of *gan* and *can* in auxiliary uses from those where this verb and its variant are associated with a full, i.e. ingressive, meaning, as reported by various sources in the literature.

However, in OE, the predecessors of auxiliary verbs in ME shared a number of properties with the rest of the class of verbs and so behaved like members of this class: a) they showed major contrasts of tense, mood, person and number, with non-finite forms being recorded for some of them only; b) most of them occurred with meanings aligning them with lexical verbs; c) their subcategorisation for the plain infinitive was not as distinctive, as it was generally found with lexical verbs; and d) they appeared in inversion and with the *ne* (later *not*), both features typical of lexical verbs at the time. In addition to these characteristics, there is also diachronic evidence demonstrating further categorisation of verbs in the way they develop. According to Warner (1993: 101), “[t]his further evidence is important because it shows their categorisation as verbs was sufficient to provide a basis for historical

⁴ The preterite-present verbs in OE are: *cann* ‘to know or be acquainted with, to be able to, to know how to’ (*OED* s.v. *can* v.¹), *dearr* ‘to have boldness or courage’ (*OED* s.v. *dare* v.¹), *mæg* ‘to be strong, have power or influence, be able to, be allowed’ (*OED* s.v. *may* v.¹), *mot* ‘might, was able or permitted to, could’ (*OED* s.v. *must* v.¹), *sceal* ‘to owe, shall, ought, must’ (*OED* s.v. *shall* v.), *þearf* ‘to be under a necessity or obligation’ (*OED* s.v. *tharf, thar* v.), *uton* ‘to let us’ (*OED* s.v. *ute* v.) and *wile* ‘to desire, wish for, have a mind to, want, intend’ (*OED* s.v. *will* v.¹).

⁵ Roberts & Roussou (2003: 36-37) list a number of characteristics of modal (auxiliary) verbs, or modals, in Modern English. According to them, modals: a) lack non-finite forms; b) cannot be iterated, except in Scots, Northern English and southern dialects of American English, where double and, even, triple modal constructions are found; c) occur with the plain infinitive and lack complements of all other types; d) are in complementary distribution with the auxiliary verb *do* and always occur before the negative particle *not*; e) always move to the complementizer position in inversions; f) can license fronting, unlike lexical verbs, in addition to ellipsis; and g) can phonologically contract, unlike lexical verbs.

change.” It also demonstrates that verbal status was maintained in ME, although to a varying degree; for example: a) some auxiliary verbs occur in ME with a present indicative ending formed on the model of lexical verbs, although they also have no third person singular ending; b) not all modal verbs within an auxiliary group occur in non-finite constructions. Warner (1993: 145), for example, shows that in ME, only *mot*, *must*, *shall* and *þearf* do not have non-finite forms⁶; however, evidence shows that non-finite constructions of the successors of other modal (auxiliary) verbs in OE continue to be present in ME, although they have no restriction to the plain infinitive and can occur with the *to*-infinitive (Warner 1993: 147); and c) some auxiliary verbs in ME occur in impersonal constructions, whereby they take an oblique NP and a plain infinitive; but, this development is shared by lexical verbs and the ancestors of auxiliary verbs (Warner 1993: 126 finds that in ME, *may*, *shall* and *will* are well attested, with *mot* and *mun* being only less central to occur in interpersonal constructions. However, he finds no instances of *can*, *dare* and *uton* in this function, presumably because these verbs select their subjects, and so “they cannot be transparent to the semantics of the construction that they enter”).

The grammatical status of *gan* and *can* as auxiliaries in ME and Scots is mentioned in many grammars (Mossé 1952: 103, Mustanoja 1960: 611, Visser 1969: 1571, Macafee 1992/93: §8.13, Moessner 1997: 114) and some textbooks (Denison 1993: 322, Burrow & Turville-Petre [1992] 1996: 47, Fulk 2012: 106-107).⁷ For example, Mustanoja (1960: 611) and Visser (1969: 1571) argue that this verb and its variant function as an auxiliary or periphrastic verb in ME. Denison (1993: 315, 322), on the other hand, discusses *gan* under the heading of “[m]arginal modals”, implying that this verb shares, at least some, characteristics of modal (auxiliary) verbs in Modern English (ModE). The *Middle English Dictionary (MED)* (s.vv. *can* v.; *ginnen* v. 3b)

⁶ Warner (1993: 147) observes that the central meaning of *mot*, *must*, *shall* and *þearf* is that of obligation and necessity. He then argues that for ME, “[p]reterite-present verbs subcategorized for the plain infinitive which denote necessity, obligation and related notions of futurity are finite only”. However, he points out in relation to the occurrence of *mun* in the infinitive, “*MED* is silent on the question of nonfinites of *mun*, but the only nonfinite citation is in the sense ‘be able’, so here too [the above statement about the semantics of obligation and necessity] may hold good” (Warner 1993: 147-148).

⁷ On the status of OE *onginnan* and OE and eME *aginnan* as auxiliary verbs, cf. Mustanoja (1960: 610-611), Visser (1969: 1580), Brinton (1988: 159-161), Sims (2008; 2014). On the grammatical status of OE *onginnan*, OE and eME *aginnan*, as well as OE *beginnan* in Ælfric, cf. Los (2000; 2005: 88-99).

refers to **gan** as “a weak auxiliary used with infinitives” and **can** as “[a] modal verb”, whereas the *Oxford English Dictionary (OED)* (s.vv. *can* v.²; *gin* v.¹) sees the former as a “mere auxiliary”, and the latter as a “periphrastic auxiliary”. Finally, van Gelderen (2004: 281) refers to **gan** as “semi-auxiliary”, although in passing only. These statements outlining the status of **gan** and **can** in ME and Scots echo the fact that this verb and its variant do not, in fact, fully fit into the characteristics of auxiliaryhood in ME identified in the preceding paragraphs. For instance, Moessner (1997: 114) argues about **gan** in Older Scots (OSc), “[a]n expansion by *gin* (present) or *gan* (past) can be followed by the bare infinitive or by the *to*-infinitive. Although formally similar to the expansion by a modal auxiliary, it is syntactically and semantically different; syntactically, because it does not combine with other expansions, semantically, because *gin/gan* does not add a modal component.” Similar statements can be found in the literature about the complementation and meaning of **gan** in ME, too (Warner 1993: 139). Therefore, it seems that the story of this verb and its variant is more complex, and so more interesting, especially when compared to the etymologically related **beginnen**, than some accounts presented in the literature on the subject of these verbs (cf. Section 1.4). The relevant auxiliary characteristics outlined by Warner (1993) are discussed below with reference to **gan** and **can**, as well as **beginnen**.

In relation to the characteristics of some auxiliary verbs in ME having finite forms and preterite-present morphology, the verb **gan** in auxiliary use occurs in the preterite, predominantly, in ME, and, exclusively, in Scots. In a similar way to **gan**, the variant **can** also occurs in finite form only in ME and Scots (*Dictionary of the Scots Language (DSL)* s.v. *can* v. 2, *MED* s.v. *can* v., *OED* s.v. *can* v.² I 1a, Warner 1993: 144). In contrast, the etymologically related verb **beginnen** is found in a range of finite and non-finite forms in ME and Scots, typically found on lexical verbs in both varieties (*DSL* s.v. *begin* v., *MED* s.v. *biginnen* v., *OED* s.v. *begin* v.¹, Visser 1969: 1373-1375, 1580-1581, Brinton 1981: 181-186; 1983; 1988: 158-159, Macafee 1992/93: §8.11, Ogura 1997, Sims 2008: 151-162; 2014). The occurrence of **gan** and **can** in finite forms only suggests a degree of auxiliatation, but does not account for this verb and its variant occurring in the preterite only. Sims (2014: 68), however, argues

that the grammaticalization of preterite forms of verbs, before or instead of present indicative ones, is not unusual. According to her, “[v]an Gelderen (2004: 180-199), for example, shows that while present tense *see*, a ModE perception verb, does not grammaticalize, preterit *saw* (with a bare infinitive complement) does.” But, although the verb *saw* may have a function of an auxiliary which is used to denote an event or action that is complete, it is crucially the role of what Sims (2008: 140) refers to as the narrowing environment of verse and the plain infinitive that helps determine whether this verb is used as auxiliary. On the other hand, even though most sources find that *gan* predominantly occurs in the preterite when this verb is found in auxiliary use, the *MED* (s.v. *ginnen* v. 3b) and the *OED* (s.v. *gin* v.¹ 1a) also list instances of this verb in the present indicative in auxiliary use in ME, whereas Moessner (1997: 114) refers to “*gin* (present)”, in addition to “*gan* (past)” in Scots. (Although non-finite forms are reported for *ginnen* in absolute and transitive uses, the reader is referred to Section 1.5.1 on the treatment of *gan* and *ginnen* in this study).

In addition to being mainly or exclusively used in the preterite, the verb *gan* is considered to be morphologically fixed, as part of the grammaticalization process (*DSL* s.v. *can* v. 2, *MED* s.v. *can* v., *OED* s.v. *can* v.² I 1a). For example, both Denison (1993: 322) and Brinton (1996: 81) refer to *gan* as being “virtually fixed” and “[morphologically] fixed in the third-person preterite form, most often singular”, respectively (cf. also *DSL* s.v. *gan* p.t., *MED* s.v. *ginnen* v., *OED* s.v. *gin* v.¹, Mossé 1952: 103, Brinton 1981: 156; 1988: 119-120, 161, Ogura 1997: 403 and Sims 2008: 150; 2014: 67 amongst others). Warner (1993: 144), on the other hand, describes the variant *can* as a new development in ME, alongside the northern and midland *mun* ‘shall, must’, observing that it displays preterite-present morphology in ME when having “‘modal’ (or ‘auxiliary’) senses”. However, although *gan* and *can* are fixed in the third person singular and have finite forms only, plural forms of this verb and its variant are found, too. For example, Brinton (1996: 81) alludes the existence of forms of *gan* other than in the third person singular in ME, especially when she lists all forms of this verb in her appendix, whereas the *MED* (s.vv. *can* v.; *ginnen* v.), the *OED* (s.vv. *can* v.²; *gin* v.¹) and Denison (1993: 322) explicitly refer to plural forms

of this verb and its variant in Southern Middle English (SME). However, the occurrence of *gan* and *can* in singular and plural forms in the preterite alone does not suggest that this verb and its variant do not undergo auxiliation. Even the OE predecessors of modal (auxiliary) verbs in ME, especially at the beginning of the period, were inflected, i.e. *þu cannst*, *we cunnan*, *we cuðon*, except in the third person singular in the present tense (Lightfoot 2006: 94, cf. Warner 1993: 101). Furthermore, the attrition of marking of inflectional endings to differentiate between the singular and plural forms in the preterite took longer in the southern dialects of ME than in the northern ones, as discussed in Section 1.3.1. In morphological terms, therefore, *gan* and *can* appear to meet specific criteria for auxiliaryhood to a greater extent than *beginnen*; the former two occur in finite forms only, whereas the latter is marked for mood, number, person and tense, although to a varying degree and differently in ME and Scots (for example, no singular and plural dichotomy is present in Scots) (*DSL* s.v. *begin* v., *MED* s.v. *biginnen* v., *OED* s.v. *begin* v.¹, Mustanoja 1960: 610-615, Kerkhof 1966: 154, Visser 1969: 1373-1375, Brinton 1981: 181-186; 1983: 241-242; 1988: 118-119, 158-159; 1996: 67, Macafee 1992/93: §8.11, Ogura 1997, Fulk 2012: 72-81).

Interestingly, there have also been further changes in relation to the morphology of the variant *can* and the verb *beginnen*; however, these take place mainly or exclusively in Scots. An example of such a further change involves the development of *couth* as a preterite form of *can* as a variant of *gan*. Unlike *gan* and *can*, this variant is found with non-ingressive meaning only, the same as characterising the variant *can*; it also exclusively occurs with the plain infinitive (*DSL* s.v. *can* v. 2, *OED* s.v. *can* v.² II, Taylor 1917, Macafee 1992/93: §8.13, J. Smith 1996: 137 and Markus 1997 amongst others). The *OED* (s.v. *can* v.²) argues that to remove the irregularity of the variant *can* appearing to be in the present indicative in form, especially when considered alongside the modal (auxiliary) verb *can*, but referring to a past event or situation, the variant *couth* was used (cf. Macafee 1992/93: §8.13). The presence of *couth* could also be indicative of further grammaticalization of *can* towards auxiliation, rather than error, especially that the *OED* (s.v. *can* v.²) also maintains that the variant *can* was probably seen as a special use of the modal

(auxiliary) verb *can*;⁸ (although the *OED* s.v. *can* v.² argues that *can* and *couth* were “[c]hiefly” used in Scots, Visser 1969: 1579-1580 has instances of this variant in NME and SME works, too, cf. Section 1.3.1). The other development concerns the emergence of *begouth* as a preterite form of *beginnen*. Like its present indicative form, this variant occurs in an ingressive sense, and with a verb in the *to*-infinitive, rarely with a verb in the plain infinitive (*DSL* s.v. *begin* v. 1). Found in Scots only, it is formed owing to association of this variant with *couth* as a preterite form of *can* as a variant of *gan*. The *OED* (s.v. *begin* v.¹) claims, “[t]he Scots forms *begouth*, *begoud*, seem due to some form-association with *couth*, *could*, probably through the

⁸ In addition to the central meanings associated with modal (auxiliary) verbs, namely deontic and epistemic, the other use of the modal (auxiliary) verbs *can* and *will* is concerned with dynamic interpretations in ModE (*dare* is unique amongst modal (auxiliary) verbs in that it has only a dynamic interpretation) (Huddleston & Pullum 2005: 55). Huddleston & Pullum (2002: 178) define dynamic modality as “concerned with properties and dispositions of persons, etc., referred to in the clause, especially by the subject NP”. They also maintain that the dynamic ability is “less central to modality than deontic permission in that it does not involve the speaker’s attitude to the factuality or actualisation of the situation” (Huddleston & Pullum 2002: 179). Examples of the dynamic *can* and *will* are in (1) below and (2) below, respectively. In the former, the focus remains on the ability of the person to beat everyone else in the club, whereas in the latter, the emphasis lies on the unwillingness or refusal on Jill’s part to sign the form.

- (1) *She can easily beat everyone else in the club* (Huddleston & Pullum 2002: 178)
 (2) *Jill won’t sign the form* (Huddleston & Pullum 2002: 192)

Although the meanings of modal (auxiliary) verbs generally represent those that are to be part of a typological account of modality, it is clear from the above descriptions that dynamic interpretations do not appear to be strictly related to modality (Denison 1993: 295, Palmer [1986] 2001: 8, 9-10, 76, 86, Huddleston & Pullum 2005: 55, Narrog 2005a; 2005b). What is more, some argue that they should even be excluded from it altogether (Steele 1975: 38, Palmer 1979: 36-37; 1986: 102-103, Bache & Davidsen-Nielsen 1997: 325, Høye 1997: 44, Gisborne 2007: 59).

In addition to meaning, dynamic modals show some differences in form, too. The dynamic *can* is marked for tense, with *could* functioning as the form referring to the past time in semantic terms, too. This is illustrated in (3) below, where the focus is on the ability of a person to play the piano in the past, rather than the possibility. Furthermore, this example illustrates the use of the dynamic *could*, with Gisborne (2007: 52) pointing out that without the parentheses, this instance would offer an ambiguous interpretation, that between a deontic and dynamic one.

- (3) *I could play the piano (until I broke my thumb)* (Gisborne 2007: 52)

The occurrence of the dynamic *could* with a reference to the past time in semantic terms is in contrast with the behaviour of modal (auxiliary) verbs expressing epistemic or deontic modality: *could*, *should* and *would* are preterite forms in relation to the sequence of tense behaviour, but in semantic terms, they do not behave like past tense forms (Gisborne 2007: 52). Gisborne (2007: 58) concludes that the dynamic *can* is like a modal (auxiliary) verb in form (finite, prototypically auxiliary in terms of negation and on formation), but in terms of semantics, it has no prototypical modal sense. This conclusion could also be applied to *can* as a variant of *gan* and *couth* as a preterite form of *can*, whereby both variants are similar in form to the modal (auxiliary) verb *can* and its preterite form *could* with deontic and epistemic meanings, except that they do not add a modal component when their meaning is concerned.

aphetic form *gan*, which became in Scots *can*, and was thus identical in form with *can* ‘to be able’” (cf. *DSL* s.v. *begin* v., Macafee 1992/93: §8.11). Interestingly, however, J. Smith (1996: 137) only lists “*bigouth, couth* ‘began’” for OSc, while “*gan, can*” for NME amongst linguistic features distinguishing between both varieties.

According to Warner (1993: 103), another characteristic of some verbs classed as auxiliary verbs is their restriction to the plain infinitive. However, there appear to be some variation in terms of the choice of infinitive with which *gan* and *can*, as well as *beginnen* occur in ME and Scots, this variation differing with respect to the verb and variety (*DSL* s.vv. *begin* v.; *can* v. 2; *gan* p.t., *MED* s.vv. *beginnen* v.; *can* v.; *ginnen* v. 3b, *OED* s.vv. *begin* v.¹; *can* v.² I 1a; *gin* v.¹ 1a, Beschorner 1920: 9-11, Mustanoja 1960: 531, Visser 1969: 1373-1375, 1379-1380, 1571-1576, Brinton 1981: 156; 1988: 120, Moessner 1997: 114, Ogura 1997: 409, Sims 2008: 148-150).⁹ *Gan* usually occurs with the plain infinitive in ME, as in (1) below, but instances of this verb have been found with other types of infinitive there. For example, Brinton (1981: 156) claims that “[*gan*] may take the simple infinitive or less frequently the *to*-infinitive or the *for-to*-infinitive”, with Warner (1993: 139) pointing out that “[in] meaningless ‘auxiliary’ uses ... even here there is variation between the *to*-infinitive and the plain infinitive [with *gan*]”. Furthermore, some sources record instances of this verb with a form in *ende* or *ing* (Visser 1969: 1379, Sims 2008: 150), while instances of *ginnen* in finite and non-finite uses, but with an ingressive meaning (*MED* s.v. *ginnen* v. 1-2, *OED* s.v. *gin* v.¹ 2-3) (chiefly non-auxiliary uses, although cf. *MED* s.v. *ginnen* v. 3b, which lists auxiliary uses of this variant, too). In Scots, the verb *gan* is only used with either the plain infinitive or the *to*-infinitive (*DSL* s.v. *gan* p.t.). The variant *can*, on the other hand, almost exclusively occurs with the plain infinitive in ME, as in (2) below (*MED* s.v. *can* v., *OED* s.v. *can* v.² I 1a), although the *OED* provides some instances of this variant with the *to*-infinitive in NME

⁹ The *OED* (s.v. *gin* v.¹ 2-3) also records instances of the verb *gan* in transitive and absolute uses in ME, with no such instances being recorded in Scots (*DSL* s.v. *gan* p.t.). In these uses, this verb has an ingressive meaning: ‘to begin, commence; to have or make a beginning’ or ‘to begin speaking, to speak. *rare*’ (absolute) and ‘to begin (something)’ (transitive). However, no instances of the variant *can* are recorded in such uses in both ME and Scots (*DSL* s.v. *can* v. 2, *OED* s.v. *can* v.²).

(*Cursor Mundi*) and Scots (Blind Harry's *Wallace*)¹⁰ only, but with ingressive meaning. In contrast, the variant *couth*, typically found in Scots, is exclusively recorded with the plain infinitive (*DSL* s.v. *can* v. 2b; *OED* s.v. *can* v.² II 4; Visser 1969: 1579), even amongst NME and SME examples (Visser 1969: 1579-1580). The verb *beginnen* is used with all types of infinitive, most frequently with the *to*-infinitive and the *for-to* infinitive in ME and almost exclusively appears with the *to*-infinitive ("rarely with the ellipse of *to*") in Scots, in addition to instances in absolute and transitive uses also found in both varieties (Visser 1969: 1373-1375, Brinton 1981: 181-182; 1988: 118-119, Ogura 1997, Sims 2008: 152). The variant *begouth*, however, is found with the *to*-infinitive only (*DSL* s.v. *begin* v. 1).

- (1) *Horn* ... *ferde to wude for to schete*;
Horn ... went to wood for to shoot;
A knave he gan imete
A knave he did meet
 (*King Horn* 940; Mustanoja 1960: 611, as cited in Brinton 1981: 171)¹¹

'Horn went to the woods to shoot; he met a peasant'¹²

- (2) *At þe last scho con hym kesse*,/
At the last she began him kiss,/
Hir leue fayre con fonge/
Her love fair did seize/
 & *went hire waye I-wysse*
 & *went her way certainly*
 (*Gawain* 1555-1557; Funke 1922: 24, as cited in Brinton 1981: 172; translation in the original)

'At the last she began to kiss him; her fair love he seized (NOT began to seize) and went her way, certainly'

¹⁰ The MS of *Cursor Mundi* quoted in the *OED* to have the line containing the variant *can* with the *to*-infinitive is Cotton Vespasian A.3 MS, whose dialect is described in *An Electronic Version of A Linguistic Atlas of Late Mediaeval English (eLALME)* as "YWR [West Riding of Yorkshire]" (*eLALME* LP18), whereas that of Blind Harry's *Wallace* is Advocates MS 19.2.2 written in a dialect "clearly ... as Scots" (Head 1997: §3.48).

¹¹ When translating examples, I provide a word-for-word gloss and a non-literal ModE translation, wherever possible. I use the term 'glossing' in a loose sense, whereby I substitute ME and Scots vocabulary (lexical glossing) or a construction (syntactic glossing) with an exact or near exact expression. All emphasis, glosses and translations are mine, unless stated otherwise.

¹² Brinton (1981: 171) translates the above as 'Horn went to the woods to shoot; a knave he met (NOT began to meet)'; I have provided an alternative translation.

Warner (1993: 110-134) also identifies the properties of auxiliary verbs at sentence-level syntax in OE and ME, such as the occurrence of such verbs in post-verbal ellipsis and in pseudogapping. However, in relation to *gan* and *can*, he argues that although this verb and its variant display this kind of property, it is only when they function as intransitive verbs (cf. Antipova 1963, M. H. Brown 1970, who report no instances of *gan* with ellipsis),

OE *onginnan* and *aginnan* ‘begin’, which might be thought to belong with a potential auxiliary group in other respects, both have an intransitive sense ‘undertake an action, proceed to an action’ which occurs in potential ellipsis contexts and elsewhere, so that no distinct construction need be involved in potential ellipsis contexts. The same holds for ME *ginnen* and derivatives. In neither case have I noted examples with pseudogapping which might support the case for ellipsis. (Warner 1993: 117)

The other property identified by Warner (1993) to be indicative of auxiliaryhood is the occurrence of auxiliary verbs as ‘intervening’ (Warner refers to such verbs as ‘I-verbs’) in impersonal constructions – they occur within such constructions, lacking a nominative subject when their dependent non-finite is impersonal. In other words, “the impersonal retains its constructional characteristics despite occurring in apparent subordination to another verb, and this other verb is in some sense independent of or transparent to the construction with which it cooccurs [sic]” (Warner 1993: 123). With reference to *gan* and *can*, as well as *beginnen*, Warner (1993: 127) claims that there are instances of *gan* in such constructions in ME, although he quotes only one, as in (3) below (and a number of instances with *onginnan* from OE). He also provides “a possible Middle English example with *biginnen* followed by a *to*-infinitive”, as in (4) below, but also observes that “*begin* to at least sometimes showed this transparency [to impersonal constructions]; here it is used, as *ginnen* often is, to put the infinitive in rhyme and probably with weakened inceptive meaning” (Warner 1993: 127, 128). However, in these examples, both *gan* and *beginnen* are used with the *to*-infinitive, rather than the plain infinitive, failing the plain infinitive criterion for auxiliaryhood, as outlined earlier. This suggests that in impersonal constructions like these ones, both verbs align with lexical verbs, rather than auxiliary ones. Overall, though, Warner (1993: 125, 129, 132-134) does not appear to regard the impersonal construction to be a strong enough indication of auxiliaryhood. But, he notes that *gan* clearly shows evidence of transparency to

impersonal in ME, a little in relation to *beginnen*, too, pointing towards the lack of distinction between impersonal and raising constructions (cf. Sims 2008: 87-95; 2014: 65-66).¹³

- (3) *He ... gan to sike*
 He ... began to sigh
Wiþ Kyng Richard gan hym euyl lyke
 With King Richard began him evil please
 (a1450-1509 *Der mittelenglische Versroman liber Richard Lowenherz* 4800; Warner 1993: 127; translation in the original)

‘He ... began to sigh, he began to be very displeased with King Richard’

- (4) *Po bigan ham alle to agrise*
 Then began them all to dread
 (c1330 *Otuel* 1604; Warner 1993: 128; translation in the original)

‘[When the Saracens saw that Otuel had killed Karnifees] then they all began to be terrified’

Finally, even though Warner (1993) does not mention changes in meaning as one of the criteria for auxiliaryhood, both *gan* and *can* occur in ME and Scots with a non-ingressive meaning, with Fulk (2012: 106) referring to the former verb as “a colorless auxiliary equivalent to ‘did’”¹⁴ (cf. also *DSL* s.vv. *can* v. 2; *gan* p.t., *MED* s.vv. *can* v.; *ginnen* v. 3b, *OED* s.vv. *can* v.² I 1a; *gin* v.¹ 1a, Taylor 1917, Funke 1922: 14, Koziol 1932: 133, Mustanoja 1960: 611, Kerkhof 1966: 156-158, Visser 1969: 1379, 1571-1575, 1577-1579, Brinton 1981: 159, 161; 1983: 235, 236; 1988:

¹³ Los (1998: 15) argues that OE *onginnan* is not usually used in impersonal constructions, always requiring a nominative subject. Although she provides one instance of this kind of construction from Denison (1993: 148), she observes a tendency for transparency to the argument structure of the impersonal infinitive *langian* ‘grieve’. She argues that transparency is only found with modal (auxiliary) verbs in OE. This transparency also characterises other instances of *onginnan* with the plain infinitive, but not the *to*-infinitive.

¹⁴ Funke (1922: 15-16), Mustanoja (1960: 602) and Visser (1969: 1572) suggest that the verb *gan* and its variant *can* bear a resemblance in meaning to the verb *do*, while the *MED* (s.v. *ginnen* v. 3b) shows that these verbs alternate in different MSS of *Cursor Mundi* (the distribution of the verb *do* in northern dialects of English is problematic, too – only isolated instances of causative *do* and periphrastic *do* are recorded in northern verse, but not in prose, of the 15th century, as discussed by Nurmi (1999: 233-234). Suggestions have also been made in the literature that the verb *gan* may be similar to *gar* in terms of the distribution (Ogura 1997: 425; 1998: 300). However, the verb *ger* has the following similar, but unrelated, meanings to *gan*: ‘to prepare or equip (oneself)’, ‘to make (sth.), to cause (sth.)’, ‘as auxiliary or quasi-auxiliary: with inf. to cause (sth. to be done) (sb. to do sth.)’ (*MED* s.v. *gēren* v. 1-3).

120, Denison 1993: 322 and Ogura 1998: 300 amongst others), and the *MED* (s.v. *ginnen* v. 3b) pointing out that the meaning of *gan* can be similar to that of *do* in the present indicative. Although the use of terms like ‘auxiliary’ in reference to *gan* and *can* suggests a loss of lexical meaning, these sources also acknowledge that instances of this verb and its variant are used with ingressive meaning, with Visser (1969: 1573) pointing out that “[n]ot in all of them is the ingressive meaning of *gan* = *commence* demonstrably non-existent”. More specifically, Visser (1969: 1379) also argues that the distinction between the meanings of *gan* and *can* is “largely arbitrary” (cf. Einkenkel 1891, Taylor 1917, Kerkhof 1966: 158-159, Brinton 1981: 161, Denison 1993: 322). The verb *beginnen*, on the other hand, has an ingressive meaning, but instances of this verb with non-ingressive meaning are also very occasionally found in ME (Visser 1969: 1373-1375, Brinton 1981: 181-182; 1988: 118-119, Ogura 1997, Sims 2008: 152). For example, the *MED* (s.v. *biginnen* v. 6) provides the following definition of this meaning of this verb, “emphasizing the occurrence of an event or act denoted by the following infin.: e.g. **beginneth springe**, does leap, **began to calle**, did call [Cp. the frequent use of *ginnen*, *gan* with this function.]’. Examples of *beginnen* with ingressive meaning and with the *to*-infinitive in ME and Scots are provided in (5) below and (6) below, respectively. An example of this verb used in a non-ingressive sense is provided in (7) below.

- (5) *Vor huanne þou begonne libbe: anhaste*
 For when you began live: immediately
þou begonne to sterve
 you began to die
 (Ayenbite p. 71, 25; Brinton 1981: 181-182; translation in the original)

‘for when you began to live, immediately you began to die’

- (6) *Sene we begyne to speik of the state of*
 Then we begin to speak of the state of
innocens;
 innocence;
 (1490 De Irlandia *The Meroure of Wyssdome* 60/6; *DSL* s.v. *begin* v. 1a)

‘Then we begin to speak of the state of innocence;’

- (7) *Þe wynt bi-gon þe schip to driue, til þei*
 The wind began the ship to drive, until they

bi-gonne *to* *aryue* ... *In* *rome*
 did to arrive ... in Rome
 (St. Alex. [I] [Vrn 46.241; Visser 1969: 1581, as cited in Brinton
 1988: 159; translation in the original)

‘the wind began to drive the ship until they arrived (NOT began to arrive) in Rome’

From this overview, it appears that the verb **gan** and its variant **can** are more advanced in the divergence of their development towards auxiliation than the verb **beginnen**, all showing variation in terms of form and the subcategorisation of complement (the potential restriction of **gan** and **can** to finite forms is discussed in Section 1.5.1). Furthermore, the extent of the divergence of these verbs seems to be more prominent in Scots than in ME, i.e. there are further developments, such as **couth** as a preterite form of **can** as a variant of **gan** recorded mainly in Scots, but also in other varieties in ME, and **begouth** exclusively found in Scots. Finally, references to **gan** and **can**, as well as **beginnen** are made with respect to ME and, occasionally, to Scots, with no statements identifying differences and/or similarities between these verbs in NME and ESc. Therefore, using late 14th and 15th century Northern English and Scots textual evidence from *The Helsinki Corpus of English Texts (HCET)*, *The Innsbruck Corpus of Middle English Prose (Sampler) (ICoMEP)*, *Teaching Association for Medieval Studies (Middle English Text Series Texts Online) (TEAMS)*, *The Helsinki Corpus of Older Scots (HCOS)* and *A Linguistic Atlas of Older Scots (LAOS)*, I investigate in this study whether the verb **gan** and its variant **can**, as well as the verb **beginnen** develop differently in terms of grammaticalization in NME and ESc. I address the following questions in this thesis:

- a) Do **gan** and **can**, as well as **beginnen** differ with respect to their morphological paradigms, in view of what we know about grammaticalization and the development of invariant forms?
- b) Do these verbs differ with respect to their complements, in view of claims in the literature that the more grammaticalized variant takes the plain infinitive; and

- c) Whether *gan* and *can* are a development from OE *onginnan* and OE and eME *aginnan*, originally expressing ingression, but shown in the literature to have undergone semantic bleaching in OE and in eME period?¹⁵

1.3 Northern Middle English and Early Scots as Related Varieties

1.3.1 Linguistic Divergence within the Northumbrian *Sprachgebiet*

The relationship between both NME and ESc is particularly striking. With Scots being descendant from the Northumbrian dialect of OE (Macafee n.d.: §1.1.2), the received view is that the labels ‘NME’ and ‘ESc’ are used to refer to a common speech area, or *Sprachgebiet* (Williamson 2002: 253). For example, in his work on the pronunciation, grammar, and historical relations in the dialect of the southern counties of Scotland, Murray (1873: 29; italics in the original) refers to “*Early Lowland Scotch*” as “the northern English ... spoken from the Trent and Humber to the Moray Forth [sic]”. In other sources, because of the linguistic similarity of NME to Scots, both varieties are treated as if belonging to one common speech area, too. For example, Horobin & J. Smith (2002: 117) discuss “[a] peculiarity of Northern ME – and Older Scots – ... ‘Northern Personal Pronoun Rule’” under the heading ‘Northern’, making a formal distinction between both varieties. However, the inflectional endings quoted by them are found in both NME and Scots, although exclusively in the latter variety, and the example provided to illustrate the above grammatical pattern in “Northern ME” is provided by them from Barbour’s *The Bruce*, an ESc text, rather than an NME one. King (1997: 158) outlines that “‘Northern’” in her work will refer to “Lowland Scots and Northern English”,

¹⁵ The following are excluded from the analysis in this study, since they do not occur in the NME and ESc data sets used here, although I still refer to them throughout this work, as and when required: a) *onginnan* is common in OE, but does not survive beyond the eME period (*OED* s.v. *ongin* v.) – Visser (1969: 1372) has one instance of this verb from as late as c1200 (*Ormulum*), and both the *MED* (s.v. *aginnen* v.) and the *OED* (s.v. *ongin* v.) have one from c1225 (*St. Juliana of Cumae* written in a dialect of SME); and b) *aginnan* is recorded in OE and in eME (*OED* s.v. *agin* v.) – Visser (1969: 1372) has instances of this verb in SME until the 14th century; in relation to texts from the north, Sims (2008: 105; 2014: 64) finds no instances of this verb in the “Northern and Mercian *Matthew*”, i.e. the Northumbrian (*Lindisfarne Gospels*) and Northumbrian/Mercian (*Rushworth Gospels*) versions of *The Gospel of Saint Matthew* in OE, although Ogura (1997: 415) has one instance of this verb in the Northumbrian (*Lindisfarne Gospels*) version of *The Gospel of Saint Luke*, in addition to one instance of *onginnan*. Nevertheless, despite some evidence of this verb being recorded in the northern varieties of OE and eME, reference sources like the *DSL* (s.v. *gan* p.t.), the *MED* (s.v. *aginnen* v.) and the *OED* (s.v. *agin* v.) do not appear to have a record of this verb in NME or ESc. For the treatment of ME *ginnen*, see Section 1.5.1.

implying that there is no linguistic difference between both varieties. Furthermore, this perception of one common speech area not being linguistically diverse is also present in descriptions dealing with the distribution of *gan* and *can*. For example, Taylor (1917: 573) focusses on the use of *can* and *couth* in Scots verse, but like Brinton (1981: 158), appears to use the terms ‘northern’ and ‘Scottish’ interchangeably.

Even though the sources mentioned in this section refer to the varieties used in the north of England and in Scotland using the collective term ‘northern’, both NME and ESc stand for a collection of linguistic features that can be perceived as distinguishing one variety from the other (in addition to the geographical area) (Williamson 2002: 253, Kopaczyk 2013: 252). For example, the *DSL* argues that “[t]he grammar of [Older Scots] is substantially that of [NME]. It shares many developments that arose from the contact between OE and [Old Norse]”, whereas accounts on ME by Fisiak (1996: 11, 12) and Fulk (2012: 24, 26, 43, 53) refer to the north, but it is clear from these sources that Scots is treated as a separate “dialect”. The importance of linguistic differences between NME and ESc can also be observed in relation to the respective grammatical systems of these varieties. In relation to inflectional morphology, King (1997: 158) claims that the linguistic restructuring undergone by English during a change from OE to ME to EModE often consisted of the same processes, such as analogy, extension or levelling. This restructuring sometimes had the same outcome in Scots and SME. An example of an outcome that is often quoted in the literature to be divergent and typical of Scots includes the development of the variant *can* as a preterite form of *gan*, but also the variant *couth* as the preterite form of *can*, both showing some associations with the present and preterite forms of the modal (auxiliary) verb *can* (*DSL* s.v. *can* v. 2). The *OED* (s.v. *can* v.²) argues about the variant *can*,

[i]n Older Scots, forms of the past tense of *can* [as a preterite form of *gan*] like *couth*, *cowd*, *could* ... frequently replaced the original *can* [i.e. the variant of *gan*] ... in its function as a periphrastic auxiliary of the past tense, in order to remove an apparent irregularity where what was perceived as a present stem of *can* [as a modal (auxiliary) verb] ... was used as a past tense in the periphrastic construction

Although sources like the *OED* (s.v. *can* v.²) identify the development of *can* as “[c]hiefly [Scots] in later use” and that of *couth* as a feature of “Older Scots”, as shown in the above quotation (cf. J. Wright 1898: 228), Visser (1969: 1571-1575, 1579-1580) provides instances of both variants from NME texts like *Cursor Mundi* (Göttingen MS Theol. 107), as in (8) below, and SME ones like Chaucer’s *Canterbury Tales*, as in (9) below. In relation to *couth*, he, therefore, shows that although this variant was “a favourite among Scottish poets”, the adjoined evidence in his work shows that it was still used in other dialects of ME, although, on the basis of the number of his examples, perhaps, to a lesser degree.

- (8) *Bot fra þe leones him saw*
 But from the lioness him saw
wel þai couth þaire laurd knaw
 well they did their lord know
 (13.. *Cursor Mundi* (Göttingen MS Theol. 107) 12343; Visser 1969: 1579)

‘But [away] from the lioness [they] saw him, they knew their lord well’

- (9) *Wel koude he knowe a draughte of Londoun*
 Well did he know a draught of London
ale
 ale
 (c1386 Chaucer *Canterbury Tales Prologue* A.382; Visser 1969: 1579)

‘He knew well [how to recognise] a draught of London ale’

However, more relevantly to the subject matter of this thesis, King (1997: 158) argues that “the Scots and/or Northern linguistic input often differed from that of Midland, Southern and South-eastern dialects of early Middle English; and, moreover, divergent developments occurred in Scots, so that features specifically characteristic of Scots emerged as the outcome.” An example of a divergent development exclusively characteristic of Scots is the emergence of *begouth* in this variety on the parallel of *couth* (*DSL* s.v. *begin* v., *OED* s.v. *begin* v.¹). For example, Boucher (1832: lxxxiii) maintains that *begouth* and *begoude* are “an irregular præter of the neuter verb *begin*”, with “[t]he formation of this tense [being] curious. The

verb *begin* seems to have been inflected in the same manner as *can*, making *begoud* (or *begouth*, with a slight and very common variation), in the same way as the other made *could*”, as in (10) below. In a similar way, J. Wright (1898: 228) argues about the origins of *begouth*, “[t]he Sc. form *begouth* (later *begoud*) is prob. due to the analogy of *couth* (*could*), pt. of *can*.” Likewise, the *OED* (s.v. *begin* v.¹) sees the emergence of *begouth* “due to some form-association with *couth*, *could*, probably through the aphetic form *gan*, which became in Scots *can*, and was thus identical in form with *can* ‘to be able to’”. Macafee (1992/93), on the other hand, only mention such developments in passing. For example, when discussing “hypothetical forms that would have resulted from the corresponding Old English forms [of Class III of strong verbs]”, Macafee (1992/93: §8.11) only lists *begoud* and *begouth* as alternative preterite forms in the singular and *begoud* as an optional past participle form, whereas sources like King (1997) do not refer to such forms of *beginnen* at all. (Other innovations typical of Scots, but not NME, are provided in Macafee n.d. and Williamson 2002 and so will not be repeated here, cf. J. Smith 2012: 45-50 on OSc characteristics).

(10) *Thai* *becouth* *to* *fle* *away*
 They began to flee away
 (1380 Barbour *Legends of the Saints*¹⁶ xxxiii 284; *DSL* s.v. *begin* v.
 1bb)

‘They began to run away’

¹⁶ There are two versions of *The Legends of the Saints*: the northern one and the southern one. Although sources like the *DSL* and Macafee (1992/93) use the generic name *The Legends of the Saints* only, I make references to *The Scottish Legendary*, unless stated otherwise, because of the practice employed by the *TEAMS*. However, there is a question of authorship of *The Scottish Legendary*. Although initially ascribed to John Barbour in all or part of a translation of the legends of the saints (Duncan 2004), his authorship has been disputed,

it seems to me to be regretted that the name of Barbour was ever associated with the legends as their author. So far as I know, there is not a tittle of evidence to prove that he had anything to do with them the facts (1) that they are written in the same dialect and in the same measure as ‘The Bruce’; (2) that the author of the first fourteen legends was an old man; (3) that he was in Holy Orders; (4) that in his old age, and when writing the legends, he was incapable from the want of health of discharging the active duties of his office; and (5) that the author of the Julian legend had “trawalyt oft in sere place” when he was a “zunge mane” ... there is absolutely nothing in these facts to prove the Barbour authorship. (Metcalf 1896: xxiii, cf. Duncan 2004)

Despite specific developments that chiefly or exclusively characterise Scots, both NME and ESc usages are also traditionally seen as grammatically innovative when compared with those found in the southern varieties of ME. An example of an innovation associated with both of these northern varieties is the development of the northern preterite during the transition from OE to ME. This change firstly occurred in the north, though sources make no distinction here between the varieties spoken in the north of England and in Scotland, and consisted of: a) the levelling of the preterite singular under the vowel of the first and third person singular, leading to a stable singular/plural opposition; and b) the removal of the number opposition at a later date, resulting in marking of the concord, if any, by endings (Wyld 1914: 205-206, Fulk 2012: 75-76, cf. Lass 1992: 130-132). This northern development operated on verbs belonging to the former Class III of OE strong verbs (OE *onginnan* and OE and eME *aginnan*, as well as OE *beginnan*; *Dictionary of Old English (DOE)* s.vv. *be-ginnan* v.; *ginnan* v.) and so, it is particularly relevant to *gan* and *can*, as well as *beginnen* in NME and ESc. A number of other innovations are seen as originating in the north. For example, in the Northumbrian dialect of OE, the present tense system was different from that in southern dialects. Lass (1992: 136) argues that it had “a highly innovative (probably Scandinavian-influenced) present, with frequent collapse of second- and third-person singular and of both with plural, and an ending in –s for all three collapsed categories.” This innovative system gradually spread from the Northumbrian dialect of OE to Scots, and then to the northern parts of the midlands during the course of the ME period, eventually reaching the south, too (Macafee n.d.: §7.8.1), although it should be remembered about the lack of written evidence in Pre-literary Scots which would allow a full lexico-grammatical analysis, as discussed below.

1.3.2 Availability of Written Evidence

In addition to both NME and Scots being seen as innovative varieties, especially when considering the changes occurring in them and those in the southern dialects of ME, it is also necessary to remember the availability of the material available before, and at the start of, the period that conventionally begins the linguistic history of NME and ESc as two separate, but similar, varieties of ‘English’. More specifically,

there is a scarcity of literary material in relation to the predecessor of the ‘English’ of the six northern counties of England before 1375, i.e. the Northumbrian dialect of OE, and a lack of this kind of material permitting lexico-grammatical analysis in relation to the ‘English’ of Scotland before 1375, namely Pre-literary Scots. With reference to the Northumbrian dialect of OE, Cuesta & Silva (2008: 50-51) provide a list of the available textual evidence in this dialect: *Lindisfarne Gospels* and *Rushworth Gospels*¹⁷, some runic inscriptions (Ruthwell Cross and the Franks Casket), three short poems (Caedmon’s *Hymn*, Bede’s *Death Song* and the *Leiden Riddle*) and some 9,000 names in Bede’s *Historia Ecclesiastica* and in *Liber Vitae Dunelmensis*, this being some, but limited, evidence from this variety. In relation to Pre-literary Scots, on the other hand, the availability of written evidence appears to be even scarcer; this is mentioned in most literature dealing with the early history of Scots. For example, both Macafee (1992/93: §1.4) and King (1997: 156) discuss the scarcity or apparent lack of literary material in this variety: the former says, “the dearth of Scots text between the early Old English period and the late fourteenth century”, whereas the latter argues, “[w]e have very little evidence of the kind of evolutionary or stylistic intralinguistic variation that must have existed and is to be found in later periods of Scots until the Older Scots period proper (c. 1375 to c. 1700) is under way”, suggesting that literary material did exist in relation to Scots before 1375, although in small quantities. Indeed, Craigie (1931: vii) outlines in the introduction to *A Dictionary of the Older Scottish Tongue (DOST)* the available evidence during the Pre-literary Scots period,

[t]he continuous Scottish record begins only with Barbour’s *Brus* in 1375, and the contemporary Legends of the Saints. Before that date the scanty evidence has mainly to be gleaned from the place-names and personal designations occurring in the Latin charters or other documents relating to the south-eastern counties, or from legal terms employed in a Latin context in the early laws. These sources have been carefully examined for this evidence, which enables the Scottish form of a considerable number of common words to be traced back to the 12th or 13th cc.

¹⁷ Although *Rushworth Gospels* is subsumed by Cuesta & Silva (2008: 51) under “Northumbrian textual evidence”, these scholars acknowledge the fact that in relation to the scribes’ dialects, Owun’s gloss of *Rushworth Gospels* is in the Northumbrian dialect of OE, as opposed to Ferman’s gloss, which is in the Mercian dialect of OE (cf. Footnote 15).

However, even though Scott (2004: 214) argues, “Craigie is perhaps a little unjust in his reference to the ‘scanty’ pre-literary evidence [because a] surprising quantity of material is available for analysis, and it can be used for much more than simply identifying older examples of common words”, the kind of linguistic evidence discussed by Craigie (1931) is of specific kind, namely toponymic, onomastic and legal. Therefore, it would be more useful in carrying out a lexico-etymological analysis, but less so when attempting to perform a lexico-grammatical one, which is what this study does.

Therefore, the lack of textual evidence in Pre-literary Scots allowing a lexico-grammatical analysis could have an effect on our understanding of the divergence of *gan* and *can*, as well as *beginnen* in NME and ESc. In particular, there may be a problem with investigating the types of changes affecting the predecessors of these verbs in Pre-literary Scots (although the inadequate evidence in the Northumbrian dialect of OE may also become an obstacle in an effort to look into the changes undergone by the predecessors of these verbs in this dialect). However, despite the dearth of written evidence in Scots before 1375, it might still be possible to attempt to reconstruct developments concerning *gan* and *can*, as well as *beginnen* there, albeit not in detail. Macafee (1992/93: §1.4), for example, argues that “Scots can more or less be identified with Northern Middle English in the early Middle English period, though it probably lagged behind somewhat with regard to those changes that originated in Anglo-Scandinavian”. This, nevertheless, suggests a much closer relationship between Scots and NME, as maintained by King (1997: 158), as outlined in Section 1.3.1. Furthermore, King (1997: 157) proposes that the range, quantity and subject matter of written evidence from 1375 onwards in Scots appears to be “impressively large”, namely, “Scots was the medium of, for instance, charters, indentures, Acts of Parliament, burgh records, the Laws of Scotland, diplomatic letters, trials, histories, treatises ... handbooks ... diaries and private letters, as well as a vast body of literature” from this time on. Nevertheless, caution should be exercised when interpreting statements about the availability of written evidence, as this “impressively large” body of texts is unlikely to be immediately available at the beginning of the ESc period.

1.4 Approaches to the Study of *Gan* and *Can*

1.4.1 Semantic and Pragmatic

Most research focussing on *gan* and *can* as auxiliaries is characterised by, mainly, a semantic-based approach, with studies largely discussing prototypical, i.e. non-ingressive, instances of this verb and its variant (cf. Wuth 1915: 55, Beschorner 1920: 13-14, Funke 1922: 1, Mustanoja 1960: 613, Kerkhof 1966: 156-158, Visser 1969: 1573-1574, 1575-1577 and Brinton 1981: 161, 170-173; 1983: 236; 1988: 117-122, 152-161 amongst others). A particular problem found in the literature on the semantics of *gan* and *can* is distinguishing non-ingressive instances of this verb and its variant from ingressive ones. Brinton (1983: 238, 240) discusses criteria for isolating non-aspectual functions of ME *gan*. According to her, under the following circumstances, an ingressive meaning of *gan* is impossible, given that the meaning of this verb interacts with the Aktionsart of the verb in infinitive and other elements within a clause in a number of ways: a) when the verb in infinitive is an achievement verb, expressing a punctual event whose inception and termination points coincide; b) when the subject or the object of the verb *gan* and a punctual (achievement) verb in infinitive is a non-plural or non-count noun, as opposed to plural count or mass nouns, which allow an iterative, and so ingressive, interpretation of *gan* and *can* in such cases; and c) when *gan* is found occurring with adverbials (or adverbs in Brinton 1983; 1988) of duration or iteration (Brinton 1981: 173; 1983: 240-241; 1988: 157-158). Others compare the meaning of this verb to that of *do* on the basis of exchangeability of *gan* and *do* in verse, as sometimes shown to be the case in some MSS (Taylor 1917: 591, Funke 1922: 26, Mustanoja 1960: 614, Visser 1969: 1572-1573, cf. Sims 2014: 67), as well as the historical present (Mustanoja 1960: 614). Sources dealing with the meaning of *gan* and *can* also highlight the arbitrariness of meaning of this verb and its variant in ME (Einenkel 1891, Wuth 1915: 55, Taylor 1917, Kerkhof 1966: 158-159, Visser 1969: 1371, Brinton 1981: 159, Denison 1993: 322).

In addition to approaching the subject of the verb *gan* and its variant *can* from a semantic-based approach, a significant amount of research has also been devoted into

different functions of this verb and its variant. As *gan* and *can* occur in narrative verse in ME and Scots, with isolated instances of the former in ME prose (Smyser 1967: 69, Terasawa 1974: 94-96, Brinton 1981: 164; 1988: 120; 1996: 69), the most frequently quoted roles of this verb and its variant are: a) to fill in a line of verse with an extra syllable, given the non-ingressive meaning associated with both *gan* and *can* (Taylor 1917: 574, Beschorner 1920: 15-19, Funke 1922: 22, Koziol 1932: 132-133, Kerkhof 1966: 154-159, Smyser 1967: 68-69, Visser 1969: 1572, Terasawa 1974: 102, Tajima 1975: 434, 437-438, Mustanoja 1983, Richardson 1991a: 80-81, Fischer 1992: 266, Ogura 1998: 300, Sims 2008: 140, Fulk 2012: 106-107, also cf. Brinton 1996: 82); and b) to make a verb in infinitive available to provide rhyme (Koziol 1932: 132-133, Kerkhof 1966: 154, 158, Smyser 1967: 74, Terasawa 1974: 102, Tajima 1975: 434, 437-438, Brinton 1981: 164). However, the use of *gan* and *can* in verse has also been linked to a specific type of text, namely rhymed narrative texts, such as romances, with the findings from Brinton (1996: 69) showing that *gan* has a “skewed” distribution in such texts, because “[this verb] seems to be associated with salient or important turns in the course of events”, “generally occur[ing] in multiples, clustering around significant events” in Chaucer’s *Troilus and Criseyde* (Brinton 1996: 77, cf. Los 2005: 92 on the skewed figures of OE *onginnan* and OE *beginnan* suggesting a discourse function of these verbs in Ælfric).

As a result of such pragmatic uses of the verb *gan* and its variant *can* in ME texts, other functions of this verb and its variant have been proposed: a) intensive-descriptive, where *gan* and *can* are used, according to Mustanoja (1960: 611-614), to provide “vivid descriptions of past events” (Funke 1922: 8, Häusermann 1930: 18-22, Homann 1954: 389-398); or b) textual, in which both Richardson (1991a: 76-89) and Brinton (1996: 67-83) show a correlation between the frequent occurrence of *gan* and the use of this verb in narrative text types in ME, as opposed to argumentative text types, whereas Los (2000; 2005: 92) demonstrates this relationship to already be the case for the predecessor of *gan* and *can*, namely OE *onginnan*, but also the etymologically related OE *beginnan*, both of which already show signs of semantic bleaching in Ælfric when used with the plain infinitive. These sources conclude that in its textual function, this verb performs various

discourse uses (foregrounding, peak marking or episode boundary marking). However, unlike Richardson (1991a), Brinton (1996: 81, 83) also argues that the verb *gan* might also acquire the interpersonal function, already referred to by Funke (1922), Häusermann (1930), Homann (1954) and Mustanoja (1960); in this function, this verb has an emphatic and intensive role, better seen as “‘internal evaluation’”, given that “*gan* is a ‘means by which the narrator makes the story interesting, highlighting the relative importance of the various narrative events’” (Silva-Corvalán 1983: 774, as cited in Brinton 1996: 81). Finally, the secondary metrical function, whereby this verb occurs in narrative verse, contributes to “[this verb’s] popularity and wide-spread use” (Brinton 1996: 72).

1.4.2 Syntactic

Apart from semantic- and functional-based approaches, other sources adopt a syntactic one to argue the divergence of *gan* and *can*, as well as *beginnen*, where relevant, along the cline of grammaticalization. However, such studies mainly concentrate on OE *onginnan* and OE *beginnan* (Los 2000; 2005), sometimes also applying the results to *gan* and *can*, as well as *beginnen* in eME (Ogura 1997; 1998; 2013, Sims 2008). For example, Sims (2008) addresses the function and structure of ingressive aspectualizers in these varieties from within the framework of generative linguistics and grammaticalization. Although she makes references to the semantics and function of OE *onginnan* and eME *gan* in verse texts, she largely explores the degree of grammaticalization of this verb and other markers by providing their structural analysis. She concludes that the syntactic behaviour of OE *onginnan* and OE *beginnan* differs in impersonal constructions with non-nominative subjects, patterns of word order and prefixed (*ge-*) complements: “the choice of complement type is influenced by the prefix (*on-* vs. *be-*) of the aspect marker, and that *to-*, when selected, heads its own projection (specifier, E) inside the VP-shell ... the syntactic structure of the non-finite complement is relevant to the structural position of the ingressive marker” (Sims 2008: iv). Another study that adopts a syntactic approach is by Los (2000; 2005: 88-99), who argues that OE *onginnan* and OE *beginnan* have different uses that are the result of the interplay of semantic and aspectual factors between the plain infinitive and the *to*-infinitive, as well as discourse effects found in

narrative contexts. She concludes that the plain infinitive is used with these verbs when they have a non-ingressive meaning, which in turn points to such constructions performing a grammatical function and so being grammaticalized.

1.4.3 Grammaticalization

The verb *gan* and its variant *can*, as well as the verb *beginnen* have also been investigated in the literature in terms of the mechanisms, parameters and/or processes of grammaticalization responsible for the divergence in the development of these verbs towards auxiliarity (Brinton 1981; 1988; 1996, Ogura 1997; 1998; 2013, Sims 2008; 2014). But, there are significant differences in the amount of descriptive and analytical detail provided; studies dealing with the grammaticalization of *gan* and *can* also focus on either ME in general or southern varieties of ME, rather than NME. For example, discussing pragmatic markers in ME and EModE from the grammaticalization and functional point of view, Brinton (1996: 81) argues that *gan* in ME is characterised by “morphological ‘fixation’”, which can be seen as evidence of “deategorialization”. She also identifies “syntactic ‘fixation’”, which is observable in the verb *gan* being more often adjacent to a verb in infinitive in later texts than in earlier ones. *Gan* is also characterised by “a loss of scope, or ‘condensation’, whereby instead of taking an infinitival complement, this verb modifies a verb; on the other hand, there is also “an increase in scope”, given that this verb performs a pragmatic function, especially that this verb “relates to larger textual and situational contexts than the purely propositional *gan*”. Next, Brinton (1996: 81) identifies “‘divergence’ or ‘form/meaning asymmetry’” in relation to *gan*, given that this verb has an ingressive function in some contexts, while textual and interpersonal in others. At the same time, she argues that there is also a case of renewal, whereby the verb *beginnen* takes over the ingressive function initially associated with *gan*, resulting in the disappearance of *gan* in this function. Finally, Brinton (1996: 81) argues that having developed textual and interpersonal functions, this verb continues in metrical function, contributing to this verb’s “popularity and wide-spread use” (72). Therefore, the textual function can be seen as a diagnostic of grammaticalization, with semantic changes that Brinton (1996) identifies in

grammaticalizing forms. In contrast, Sims (2014: 68-69) outlines the following few sentences about the process of grammaticalization of *gan*,

[f]rom grammaticalization perspective, the preterite becomes frozen construction, and the over-extended use of this form contributes to its continued bleaching and eventual loss of semantic content. That a preterit form grammaticalizes before (or instead of) a present tense form is not unique ... In addition to a narrowing tense construction, *gan*'s infinitival complement becomes more restricted, occurring mainly with the bare infinitive

However, there are also some studies of *gan* and *can* that offer an insight into this verb and its variant in Scots. Nevertheless, the focus of these works remains on issues unrelated to the process of grammaticalization and/or its mechanisms and processes. For example, Taylor (1917) concentrates on the frequency of occurrence of the morphological forms and complementation of *can* and *couth* in ESc and Middle Scots (MSc) verse, more specifically Barbour's *The Bruce*, Wynthoun's *The Original Chronicle of Scotland*, Blind Harry's *Wallace*, Holland's *The Book of the Howlat* and Henryson's *The Poems of Robert Henryson*. Macafee (1992/93: §8.13) makes references to the grammatical status and use of *gan* and *can* in Barbour's and Douglas' works, citing examples from *The Scottish Legendary* and *The Troy Book*, and from Rolland's *The Book of the Seven Sages*. Moessner (1997: 114) mentions, in no more than a few lines, the status of *gan*, its complements, meaning and function performed "mainly" in verse texts. J. Smith (1996: 137) quotes the following from Jordan (1974: 18): "*bigouth*, *couth* 'began' (analogous to *cupe* 'could') [in OSc and] *gan*, *can* [in NME]" amongst linguistic features distinguishing OSc and NME. Some ME grammars make implicit references to *gan* and *can* in Scots, but with various degrees of detail (Crook 1970: 373-374). For example, Visser (1969: 1571-1580) provides instances of this verb and its variant from Scots texts, in addition to texts from other varieties, but only refers to *couth* as "[appearing] to have been a favourite among Scottish poets", with examples of this variant from NME and SME works, too (1579). Statements in reference sources like the *MED* and the *OED* are usually in relation to *gan* and *can* in ME in general, except for the mention of the variants *can* and *couth* as chiefly Scots features, especially in the descriptions provided by the *OED*.

Therefore, this study fills the gap in our knowledge of the grammaticalization of *gan* and *can*, as well as *beginnen* in NME and ESc by offering an analysis of whether these verbs differ in relation to their morphological paradigms, complements, meaning and function in these varieties.

1.5 Scope of the Work

1.5.1 *Gan* and *Can* as Separate Lexemes or Variants of *Ginnen*

I distinguish in this study between *gan* predominantly used with the plain infinitive and in non-ingressive sense (auxiliary uses), and *ginnen* found in absolute, transitive and intransitive uses (finite and non-finite forms) and with ingressive meaning (chiefly non-auxiliary uses, although cf. *MED* s.v. *ginnen* v. 3b, which lists auxiliary uses of this variant, too). Although the majority of sources deal with *gan* only without mention of *ginnen* because of the auxiliary uses of the former (*DSL* s.v. *gan* p. t., Homann 1954, Terasawa 1974, Mustanoja 1983, Macafee 1992/93: §8.13, Sims 2008; 2014, Fulk 2012: 106-107, Ogura 2013), while others appear to distinguish between these forms (*MED* s.v. *ginnen* v., *OED* s.v. *gin* v.¹, Funke 1922, Mossé 1952: 103, Mustanoja 1960: 611, Brinton 1988: 119-122; 1996: 67-83, Fischer 1992: 265, Moessner 1997: 114 and Ogura 1997 amongst others), it is still possible to come across *gan* being referred to as *ginnen* in the literature (Brinton 1981: 154-174; 1983, Warner 1993: 139). For example, Brinton (1981: 154) refers to “ginnen plus the infinitive” throughout her study, regardless of whether she means *gan* (in the preterite) or *ginnen* (present indicative). Warner (1993: 139) argues that “Middle English *ginnen* may well show a preference for the plain infinitive, especially in weakened senses”, suggesting that both ingressive and non-ingressive instances of *ginnen* are used in this function, but also non-preterite and preterite forms. Richardson (1991a: 76-89, 117) makes references to *gan* throughout, but in a summary, he mentions “ginnen plus infinitive” while denoting the former. The *OED* (s.v. *gin* v.¹) refers to the preterite form *gan* with the plain infinitive (intransitive uses) in ME, which “was commonly used in a weakened sense, as a mere auxiliary ... serving to form a periphrastic preterite”, as in (11) below, but also to a form with the meaning of ‘to begin, followed by inf. active, with or without *to*; rarely *for to*’,

which is *ginnen* (present indicative) with the plain infinitive (intransitive uses), as in (12) below.

- (11) *To bed thai gun go*
 To bed they did go
 (c1330 *Amis and Amiloun* 1161; *OED* s.v. *gin* v.¹ 1a)

‘They went to bed’

- (12) *That so loude and sore ginneth wepe*
 That so loudly and painfully begins weep
 (c1330 *Arthour and Merlin* 1329; *OED* s.v. *gin* v.¹ 1a)

‘That begins to weep so loudly and painfully’

The *OED* also identifies finite and non-finite forms in this variety with the meanings of ‘to begin, commence; to have or make a beginning’ (absolute uses), as in (13) below, and ‘to begin (something)’ (transitive uses), as in (14) below. In contrast, the *DSL* (s.v. *gan* p.t.) only mentions the preterite form of this verb “usually” with the plain infinitive in Scots.

- (13) *Pe flode bigan to gynne, & klosed it [the island]*
 The flood did begin, & closed it
aboute
 about
 (c1330 Robert Mannyng of Brunne *Chronicle* 77; *OED* s.v. *gin* v.¹ 2a)

‘The flood started and encircled the island’

- (14) *Dauid had gunen a batayl kene*
 David had begun a battle brave
 (a1400 *Cursor Mundi* 7792 (Göttingen MS Theol. 107); *MED* s.v. *ginnen* v. 2b)

‘David had begun a brave battle’

The issue here might, therefore, be whether in the case of *gan* and *ginnen*, we are dealing with a single lexeme or more than one. This matter is important because Warner (1993) identifies the restriction of some verbs to finite forms as one of the criteria determining whether a verb displays auxiliary behaviour in OE and ME (although evidence shows that non-finite constructions of ancestors of some auxiliary

verbs in OE continued to be present in ME). In other words, the absence of non-finite forms is a striking feature of some auxiliary verbs, making it “a formal distinction from the wider class of verbs” (Warner 1993: 145). For example, in relation to a modal group within an auxiliary group, as already outlined in Section 1.2, Warner (1993: 144-148) shows that *mot*, *must*, *shall* and *þearf* in ME do not have non-finite forms, this issue not being related to frequency, but to the fact that non-finites do not appear to be restricted to the plain infinitive in ME and that these verbs shared the same semantics (cf. Footnote 6). The occurrence of the preterite form of *gan* in the third person singular, with, mainly, the plain infinitive and in a weakened sense, i.e. in auxiliary uses, in ME and, exclusively, in Scots would, therefore, suggest that this verb belongs to the class of auxiliary verbs. On the other hand, the presence of *ginnen* with full morphology, with infinitive, but also in absolute and transitive uses, and in an ingressive sense, would imply that especially in ME (*MED* s.v. *ginnen* v. 1-2, *OED* s.v. *gin* v.¹ 2-3), the verb *gan* more or less remained within the more general class of verbs. Thus, although *gan* and *ginnen* seem to remain part of the same paradigm initially, when considering their etymology, namely OE *onginnan* and OE and eME *aginnan* (*MED* s.v. *ginnen* v., *OED* s.v. *gin* v.¹), it is also evident that they are differently affected by the interplay of morphological, syntactic and semantic factors. It might, therefore, be possible to speak about a split within a lexeme (cf. G. G. Corbett 2013: 17-171) between the preterite form and the rest of the paradigm, resulting in new category membership: *gan* behaving more like an auxiliary in ME and Scots, whereas *ginnen* functioning as a lexical item, but only in ME, with no instances of it recorded in Scots.

This brings me to my next point. So far, I have referred to *can* as a variant of *gan*, but without explaining why I chose not to treat it as a separate lexeme. There are a number of reasons for this: a) the variant *can* has no distinct etymologically related OE forms of its own, with OE *onginnan* and OE and eME *aginnan*, as well as OE *beginnan* (*OED* s.v. *begin* v.¹); b) the distribution of *can* is geographically restricted, this variant being mainly used in the northern and north-Midland dialects of ME (*MED* s.v. *can* v.), but also in Scots (*DSL* s.v. *can* v. 2), with *gan* being sometimes referred to as the “the Southern counterpart of *con*” (Tajima 1975: 429); c) *can*

interchanges with **gan** in some early MSS of *Cursor Mundi* (*OED* s.v. *can* v.²I 1a), with different MSS of Barbour's *The Bruce* also alternating between **gan** and **can** (Taylor 1917: 574); and d) **can** is seen as a variant of **gan** in descriptions in the literature on the subject (cf. also Funke 1922, Mustanoja 1960: 610, Terasawa 1974: 89, Tajima 1975: 429, Visser 1969: 1571 and Brinton 1981: 158, 169; 1983: 238 amongst others), with some sources like Taylor (1917: 576) even advocating not distinguishing between **gan** and **can** at all. However, even though **can** is seen by some as a variant of **gan**, it is the only one of the two that also has a distinct preterite form **couth**, whereby both are used with the same preterite meaning, as outlined in Section 1.2 (cf. also *DSL* s.v. *can* v. 2, *OED* s.v. *can* v.² II, Taylor 1917: 573, Visser 1969: 1579 and Macafee 1992/93: §8.13 amongst others). Therefore, although it also is possible to speak about a split in relation to **gan** and **can**, it is one that does not result in a development of a new lexeme. Splits within auxiliary or periphrastic constructions are not unusual, with G. G. Corbett (2013: 181) arguing that different forms of auxiliaries can be selected; these may result in splits within a periphrastic construction, leading to situations where “the lexical verb requires auxiliary X and auxiliary X determines form Y”.

In the literature, attempts have also been made to account for the fact that **can** is a variant form of **gan**. Funke (1922: 21; translation mine) sees assimilation as a contributing factor to the distinction between this verb and its variant, “es ist mir wahrscheinlich, daß es sich um eine Assimilationserscheinung des schon unbetonten *zan* an vorhergehende stimmlose Endlaute handelt im Zusammenhang mit dem Schwund des End *e* [it seems likely to me that this is an assimilation of the already not accentuated *zan* to the preceding unvoiced endings in connection to the disappearance of the ‘e’ endings]”. Visser (1969: 1571) argues that “both *g* and *c* represent allophones of the same phoneme /*g*/”, which differ in the energy of articulation. Minkova (personal communication, 28 July 2015), on the other hand, speculates that it is possible to treat **can** as a variant of **gan** on the basis of the variable rate of aspiration: “in ME the voiced velar plosive /*g*/ [sic] in initial position, especially when followed by a nasal, frequently alternated with /*k*/ [sic]”. She argues that the rationale for the neutralisation is the fact that both consonants have a variable

rate of aspiration: [k] is more aspirated than [g]. Furthermore, she also maintains that in English Gaelic influenced regions, the velar stop is realised with more aspiration. Therefore, according to her, the underpinning of the lexico-semantic confusion of *gan* and *can* is the phonetic neutralisation of [g] ~ [k], respectively. However, while these theories may have their own merit, it seems that potentially, confusion may only arise when comparing *can* as a variant of *gan* with the singular form of *can* as a modal (auxiliary) verb. For example, the *OED* (s.v. *can* v.²) argues that the variant *can* was, in its form and construction, associated with the modal (auxiliary) *can*, “[t]he word was probably often perceived as a special use of *can*”, whereas Denison (1993: 322) suggests, “[*can* is] a possible blending of forms from GINNEN and modal CAN”.

1.5.2 Periodisation of the Data Sets

The dating behind the labels ‘NME’ and ‘ESc’ used in this study does not reflect the generally accepted periodisation for these varieties. Although ‘NME’ is by convention applied to the ‘English’ of the six northern counties of England between 1100 and 1475, whereas ‘ESc’ to that of Scotland between 1375 and 1450 (Macafee n.d.: §1.1.3), the periods referred to in this study are between 1350 and 1500 for the former, and 1375 and 1500 for the latter. It should be remembered that any dates for historical varieties are always subject to different degrees of acceptance. For example, when discussing the historical date boundaries of the ‘English’ of Scotland, McClure (1994) and J. Smith (2000) claim that all such dates are approximate, while Kopaczyk (2013) proposes an alternative periodisation altogether. Görlach (2002) and J. Corbett, McClure & Stuart-Smith (2003) conclude that, “all period boundaries in historical disciplines are open to objections”, thus suggesting a possibility for flexibility in this respect. Therefore, I took into account the following two reasons when deciding on altering the conventionally accepted period dates for NME and ESc: firstly, the organisation of texts in two corpora used in this study, namely the *HCET* and the *HCOS*, follows four sub-periods in the following increments: a) 1150-1250; b) 1250-1350; c) 1350-1420; and d) 1420-1500 (with texts being classified in the diachronic part of the corpus on the basis of “the principles of socio-historical variation analysis [with] the selection striv[ing] for a representative coverage of

language written in a specific period”, as argued by Kytö 1996: §1.1); and secondly, unless a precise date of an MS can be given, the assignment of dates by quarter centuries is followed from the *MED* (Lewis 2007: 44).

In addition to changes in the way I use the labels ‘NME’ and ‘ESc’, I decided to break up the NME and ESc data sets into two periods: a) Period I, covering, as already mentioned in Section 1.1, the time span between 1350/1375 and 1420; and b) Period II, referring to the time span between 1420 and 1500. Although applied to both data sets, this division reflects two ‘literary traditions’ practised in Scots around that time: a) the older tradition represented by Barbour’s *The Bruce*, which, according to Lynch (2007: 129), can be characterised in the following way: “[i]ts combination of epic, chronicle, and romance reflects European-wide chivalric modes of writing” (this tradition is continued in Wyntoun’s *The Original Chronicle of Scotland* and in *The Scottish Legendary*, both using *The Bruce*’s metrical format, namely octosyllabic couplets); and b) the other tradition being typically courtly and lyrical, rather than epic, represented by “[i]t flavoured allegory and dream vision to exploit the imaginative and expressive potential of fiction and replaced couplets by stanzas, prioritizing rhyme over alliteration” (Lynch 2007: 129); the first and major verse text to represent this tradition is James I’s *The Kingis Quair* (c1424), given that this work illustrates James I’s accomplishment and English models (James I was affected by English court culture) (W. H. Brown 2004). Lynch (2007: 129) continues that verse texts of the 15th century in Scotland “oscillated between these two traditions as well as between native and imported influences, yielding hybrid texts that ingeniously combine genres and styles.” In this respect, I would argue that this division lends support to applying it to the ESc data set used in this study, but at the same time to the NME data set, too, in order to be consistent.

However, some works produced in Scotland in the 1420s and afterwards might display features characteristic of Anglicisation in language, given the courtly and lyrical tradition referred to above. Sampson & Churchill ([1970] 1999: 58), therefore, list the following for the periods of the “Scottish Language”: “before 1300, Northumbrian or Early Northern English; 1300-1450, Early Scots; 1450-1620,

Middle Scots. The typical examples of Early Scots are Barbour's *The Bruce* and Wymbout's *Chronicle*; of Middle Scots the writings of Henryson, Dunbar, Douglas and Lyndsay." Although their periodisation is different from that by Macafee (n.d.: §1.1.3) and others, as outlined at the start of this section, it is worth noting that Sampson & Churchill ([1970] 1999: 58-59) also identify the following, "[f]or the sake of exactness, we may distinguish an Early Transition Scots (1420-1460), typified by *The Kingis Quair*, and *Lancelot of the Laik*; but the language of these poems represents no type, literary or spoken; it is a bookish fabrication, containing southern and pseudo-southern forms derived from Chaucer." Yet, according to Williamson (2002: 253), although the emerging 'standard' variety in England gradually started to influence the written language in the north, resulting in general erosion and disuse of local varieties, this was not the case in Scotland; here, the language "as well as evolving distinctively in speech, was developing as the written medium for literature, law and administration within a separate state" (Williamson 2002: 253). I, therefore, employ the division into two periods in this study, with a view that there might be linguistic variation, not just between NME and ESc, but also over the 125/150-year time span in question, potentially brought on in Scotland by the emerging 'standard' variety in England.

1.6 Structure of the Thesis

This thesis is organised as follows: in Chapter Two, I discuss methodological issues associated with the corpora used for the investigation of *gan* and *can*, as well as *beginnen* in the NME and ESc data sets. In Chapter Three, I deal with the morphology of these verbs, arguing that the former two are subject to decategorialization and phonological reduction, as part of the process of grammaticalization, whereas the latter is not, undergoing changes affecting the category of verbs in general at the time. I also show there that the variant *begouth* in the ESc data set arises out of analogy. The type of complementation with which *gan* and *can*, as well as *beginnen* are found in ME and Scots is discussed in Chapter Four, where I also propose that the occurrence of *gan* and *can* with the plain infinitive, and *beginnen* with other types of infinitive (in addition to zero

complementation and other types of non-infinitival complements¹⁸), in the NME and ESc data sets is due to reanalysis in both cases, with evidence of analogy being present, too, although to a varying degree. Chapter Five deals with the semantics of *gan* and *can*, as well as *beginnen* in the data sets under investigation, where I demonstrate there that not all instances of the former two are subject to semantic bleaching; the verb *beginnen*, on the other hand, is almost exclusively used with ingressive meaning. This leads me to Chapter Six, where I argue that *gan* and *can* are an integral part of verse (both octosyllabic and other types of lines) as a result of changes associated with grammaticalization: I show that this verb and its variant occur in narrative verse to make the verb in infinitive available to provide rhyme, with the other functions being textual and interpersonal, where the author of a text uses *gan* and *can* to isolate events or actions, and changes in behaviour of some characters from the rest of the narrative and to emphasise their importance. In Chapter Seven, I present a summary of the findings on *gan* and *can*, as well as *beginnen* in the NME and ESc data sets, conclude about the similarities and differences between these verbs from a dialectal point of view, and deal with the limitations of this study.

¹⁸ In this study, I refer to instances of verbs with zero complementation and other types of non-infinitival complements as lexical verbs, when juxtaposing them with auxiliary verbs, including modal (auxiliary) verbs. However, when referring to verbs grouped on the basis of the infinitival complementation, I use the term ‘independent verb’, since lexical verbs can be used with infinitive, too.

Chapter Two – Methodological Issues

2.1 Introduction

This chapter presents an overview of methodological issues associated with this study. It starts with a survey of the computerised corpora used here as the source of samples, namely the *HCET*, the *ICoMEP* and the *TEAMS* for the NME data set (Section 2.2.1) and the *HCOS*, the *LAOS* and the *TEAMS* for the ESc data set (Section 2.2.2). In descriptions of the above corpora, special attention is paid to the dialect of the samples and the MS or MSS representing the samples (whether or not such samples derive from the same or different versions of MSS, especially that the same text or sample may be written in different hands, depending on the MS used). Section 2.2.3 looks into issues associated with the representativeness of this corpus investigation: the unequal size of the data sets, the types of forms of literary language and the genre types of texts used in the data sets and the implication of this on the results of this investigation. Variation of the spelling of the items under investigation is discussed in Section 2.2.4. Section 2.3 introduces the criteria (semantic and syntactic) to determine whether *gan* and *can* are in auxiliary uses in the NME and ESc data sets. Research methods employed in this study (quantitative and qualitative) are then outlined in Section 2.4, where methodological issues related to corpus analyses carried out in this study are also discussed.

2.2 Materials

Corpus linguistics can be viewed as a research approach developed to facilitate empirical investigations of language variation and use. The application of this approach results in findings that have a wider appeal in terms of generalisability and validity. While early corpus studies might have been based on simple collections of written or transcribed texts and so not being representative (McEnery, Xiao & Tono 2006: 3), it has recently been possible to witness an increased interest in computerised material to investigate language use and change. There are advantages of using electronic corpora for language study. The most obvious is the speed and the ease with which data can be processed and manipulated (searching, selecting, sorting and formatting). For example, Beal, Corrigan & Moisl (2007: 11) argue, “relative to

paper-based corpora, electronic corpora offer well-known advantages of accessibility and amenability to fast and reliable computational analysis”, whereas Viana, Zyngier & Barnbrook (2011: 116) maintain, “[t]he availability of electronic corpora and the appropriate analysis tools have made it possible to examine language on a larger scale and see new patterns which would have been beyond the capacity of the human eye.” Furthermore, computerised corpora can also be processed and manipulated accurately and consistently, this being especially the case in relation to annotated corpora (ease of extracting information, multifunctionality of annotated material, standard reference resource). The study of historical dialectology has also reached a stage in which the NME and ESc varieties can benefit from the computerisation of a range of texts through the availability of the following computerised historical corpora: the *HCET*, the *ICoMEP* and the *TEAMS* for the NME data set and the *HCOS*, the *LAOS* and the *TEAMS* for the ESc data set. In what follows, I expand on these electronic sources used for my data by providing the necessary background to them.

2.2.1 Northern Middle English Data

The NME data came from the following corpora containing samples of continuous, non-grammatically tagged texts of various lengths: a) the *HCET* and the *ICoMEP* were used to select over 370,000 words of northern samples, the former corpus consisting of this material in both prose and verse, whereas the latter in prose only; and b) the *TEAMS* was used to extract over 20,000 words of northern samples in verse, given the exclusive presence of prose texts in the *ICoMEP* and fewer verse texts in the *HCET*. The word counts of the northern texts in the corpora used for the NME data set are given in Table 1 and have a total of 393,906 words.¹⁹ Other corpora were unsuitable for the purpose of this study for the following reasons: a) they covered either earlier or later periods in the history of English (*A Representative Corpus of Historical English Registers* covers the years between 1600 and 1999; *The Lampeter Corpus of Early Modern English Tracts* contains material from the EModE

¹⁹ The lists containing a selection of northern texts from the *HCET*, revised northern texts from the *ICoMEP* and northern texts from the *TEAMS* to carry out this study are provided in Table 58, Table 59 and Table 60, respectively, in Appendix One, together with the type of the text, as classified by the compilers of the *HCET* and *ICoMEP*, and the word count for each text, including the word count for the unedited sample in the *ICoMEP* in brackets.

period; *The Dictionary of Old English Corpus in Electronic Form* features texts from the OE period; and *The Corpus of English Dialogues* covers the years between 1560-1760); b) they contained ME material, but not NME, despite covering the time span applicable to this study (*The Corpus of Early English Correspondence* covers the period between 1400-1800); c) they contained the same or nearly the same NME material as the corpora used in this study (*Corpus of Middle English Prose and Verse (CMEPV)*, *The Penn-Helsinki Parsed Corpus of Middle English (PPCME2)*)²⁰; and d) they did not have readily available information on the dialect of samples (*CMEPV*); however, although they contain information on an MS on which a scholarly edition was based to compile such a corpus, there is no facility to identify texts from northern MSS.

Table 1

Word Counts of Northern Texts Used for the NME Data Set per Period I and Period II

	<i>HCET</i>	<i>ICoMEP</i>	<i>TEAMS</i>
Period I	26,670	82,890	-
Period II	41,095	223,071	20,180

Except for the *TEAMS*, which provided complete descriptions on the preferred MS and its dialect, as well as the historical and biographical information of the author, when known, behind the text, both the *HCET* and the *ICoMEP* operate a parameter system to describe each text, allowing the user to conduct a variation-based analysis of the quantitative results. For the purpose of this study, I therefore selected samples said to be written in NME from both corpora using the following information from each COCOA header: a) the date of the MS, rather than the date of the original; and

²⁰ For example, *Cursor Mundi* and *Liber de Diversis Medicinis* found in the *HCET* are not included in the *PPCME2*, although the former work is present in the *CMEPV*. In a similar way, Richard Rolle's *Prose Treatises and Epistles*, and *The Mirror of St. Edmund* found in the *PPCME2* are not present in the *HCET*, but are already included in the *ICoMEP*, although in varying lengths.

b) the northern dialect described as either ‘NL’ and ‘NO’²¹ (= ‘Northern’). I also collected the following information: a) whether a text on which a sample was based was classed as verse or prose; and b) text type, also including its prototypical text category, since it reflects the “continuity of the types of text represented throughout the history of English” and groups texts into broader categories, such as ‘statutory’, ‘instruction secular’, ‘instruction religious’, ‘expository’, ‘narration non-imaginative’ and ‘narration imaginative’ (Kytö 1996: §3.3.4).²² However, despite using the same parameter system, both corpora displayed some variation in coding texts. For example, the *HCET* and the *ICoMEP* did not use the same periods when specifying the parameter ‘Date of Manuscript’ from the source edition. The editorial practice used in the compilation of the former corpus is to refer to the time spans of 1350-1420 and 1420-1500; in the latter, on the other hand, the practice is to refer to 50-year periods, where for example, ‘1100+’ denotes ‘1100-1149’, but also to dates like ‘-1400+’, which denote a less specific time span, namely ‘Around 1400’, with only two northern texts in this corpus containing more specific dates under this parameter: Hilton’s *Angels’ Song* dated c1400 and *The Mirror of St. Edmund* dated c1440.

Having two different systems of classifying textual evidence according to the date of an MS would mean that an interpretation of the findings uncovered in this study would be difficult to compare across the NME data set. Therefore, I adopted the convention used in the *HCET* to group texts from the *ICoMEP*, even though the former uses longer time spans, while the latter shorter and, in some cases, even less specific ones. Furthermore, the convention used in the *HCET* is already adopted in the *HCOS*, discussed in Section 2.2.2, thus making it easier to group texts into respective periods. However, in order to categorise texts into one system, I checked the dates of MSS of the *ICoMEP* in the bibliographical data used to compile it (cf. Markus 2006: §5), but also in the *eLALME* to avoid any potential misclassification. Yet, I encountered some problems with the dating of texts in the *ICoMEP*. For

²¹ The final letter in these codings refers to the source of the dialect classification: ‘L’ = *A Linguistic Atlas of Late Mediaeval English (LALME)* or ‘O’ = source other than *LALME*.

²² In a similar way to the *HCET*, each sample contained in the *ICoMEP* is also described in terms of the “[d]iachronic prototype” of each text: “*expository/instruction, religious/instruction, secular/narrative, imaginative/narrative, non-imaginative/statutory*” (Markus 2006: 60; italics in the original).

example, the dates of the two parts of the *Alphabet of Tales* in this corpus are given as 1400+ and 1450+ in the header of the respective files with the text. Using these two dates would result in a classification of both parts of this text into Period I and Period II, even though they derive from the same MS, namely Add 25719. Therefore, after checking the dates of this MS in an edition of this MS (Johnson 1993: 26) and confirming them in the *eLALME*²³, I assigned both parts of the *Alphabet of Tales* to Period II (1420-1500) in this study. While performing similar such checks on other NME texts from the *ICoMEP*, I also discovered that in some cases, the date of an MS contained in the header in the above corpus was significantly different from the date listed in an edition to which this corpus refers as its source of data. For example, *Richard Rolle of Hampole, Yorkshire Writers* has 1300+ in the header, suggesting the range of 1300-1349, but this text should be dated “[f]irst half 15th cent.,” according to the information in the *eLALME* (LP22) on MS Rawlinson C 285. In such cases, I reorganised samples into their respective periods, according to the dates of MSS provided in the original editions used to compile the *ICoMEP*.

In addition to the issue with accurate classification of texts according to their date of MS, I also encountered a problem of inaccurate dialectal information of northern samples contained in the *ICoMEP*. For example, the bibliographical data in sources used to compile this corpus in some cases provided different information (or sometimes even no information) about the dialect in which such extracts were written. For example, even though samples from *Richard Rolle of Hampole ... and his Followers, Vol. II, Part I* (rollho2a.rtf) are described in the corpus as ‘NL’, the edition that provided the source of the sample also contains extracts from MS Reg. 17 B XVII, whose dialect is “mixed and impure [by] a West-Midland man” (Horstman 1896: 1), and from MSS Cambridge L1.i.8, fol. 201 and Cambridge Add 3042, which are described in the bibliographical source as “none of [the two] retain[ing] the original (northern) dialect” (Horstman 1895: 83), with the *eLALME* (LP425) being more specific, “NFK [“Language of W Norfolk, but with slight admixture”]” and “LIN [Lincolnshire]”, respectively. Another such an instance concerns “(Walter Hilton’s) Epistle on mixed life” (roll1hor.rtf), with the

²³ No LP number is provided for MS Add 25719 in the *eLALME*.

information in the *ICoMEP* that this extract comes from “Ms. Vernon, fol. 353”. Yet, in the edition being the source for this sample, Horstman (1895: 264) argues, “Ms. Thornton ... the only northern Ms. known; Ms. Vernon, the oldest existing Ms.”, thus suggesting that MS Vernon is the oldest, but not necessarily northern, with the *eLALME* (LP7630) confirming this by listing the dialects of this MS as “CHS [Cheshire]” and “WOR [Worcestershire]”. Given that having samples from non-northern dialects would seriously affect my results, I performed checks for every sample described as northern in the *ICoMEP* and discarded any non-northern material. This was a very time-consuming process, but the only way to avoid inaccurate results.

However, the problem of not having an accurate picture of the dialectal information of samples used for an investigation is not restricted to an electronic corpus like the *ICoMEP*. For example, while Terasawa (1974) may indirectly acknowledge the differences in the dialectal distribution and the use of *gan* and *can* in northern and Scots texts, it turns out that not all of his texts should, in fact, be categorised as northern or Scots. The reason for this is that Terasawa (1974) classifies texts according to the geographical area of composition, rather than the dialect of the MS. More specifically, *Awntyrs* is marked by Terasawa (1974: 95) as northern, but the edition of this work used by him derives from Oxford MS Douce 324 (Bodleian MS 21898), described by the editors of the *TEAMS* (n.d.: The Awntyrs off Arthur: Introduction) as non-northern, “[i]ts scribe wrote in a northwest Midlands dialect, though linguistic traces in the four surviving transcriptions locate the poem’s area of composition on the northwest border of England and Scotland; given the setting of the action in Carlisle, Cumberland seems a likely place of origin”. Another example involves *Tottenham*, also described by Terasawa (1974: 96) as northern. According to the editors of *TEAMS*, this text is found in three MSS: MS Harley 5396 (fol. 306r-310r), MS Ff.5.48 (fol. 62r-66r) and MS English 590F. As the edition of this text used by Terasawa (1974: 96) and found in French & Hale (1930) is from MS Harley 5396, the dialect of this MS is “general Midland” (*TEAMS* n.d.: The Tournament of Tottenham and The Feast of Tottenham: Introduction). Likewise, with reference to *Golagrus* or *Gologras*, the edition used by Terasawa (1974: 96) is an 1827 amended

copy of a facsimile of Chepman and Myllar's type originally published in 1508 and written in "a Middle Scots dialect" (*TEAMS* n.d.: *The Knightly Tale of Gologras and Gawain: Introduction*), therefore outwith the 1400-1500 range provided by Terasawa (1974: 96).

Finally, in addition to issues with the dating and dialect of the samples used in the *ICoMEP*, I also found that this corpus contained the same material in some parts as the *HCET*, especially in relation to Richard Rolle's prose works. Naturally, I eliminated any such samples to avoid inaccurate results. However, while I was carrying out checks on both corpora to identify duplicate extracts, I noticed that the *ICoMEP* itself contains duplicate extracts of a text, even from the same MS, albeit by different editors. For example, *The Abbey of Holy Ghost* from Lincoln Cathedral Library Thornton MS occurs in this corpus twice: the first copy is edited by Horstman (1895) (rollho1.rtf), while the other one by Perry (1867) (abbey.rtf). Different editions can differ in quality as well as the conventions used by their editors when identifying specific linguistic features – this can be seen with respect to *The Abbey of the Holy Ghost*, but with only some minor variation between both editions, this variation mainly concerning punctuation and the orthographic presence or absence of the final <e> on, for example, verbs (the lack of inflectional ending <e> and <en> to signal the plural of verbs in the preterite is to be expected, given statements, for example, by Fulk 2012: 74 about the loss of this inflection on verbs in this tense in the north, cf. J. Smith 1996: *Intralinguistic Correspondences I: The Loss of Final -E*, Duffell 2008: 96 on the loss of the final <e> in the northern varieties of ME and in Scots, respectively). I, therefore, felt that either copy of *The Abbey of the Holy Ghost* was acceptable to be used in this study, and so I included Perry's (1867) edition of this work in the NME data set. Having conducted all the additional checks, as outlined in this section, I was satisfied that I had eliminated issues that would otherwise affect the outcome of this investigation.

2.2.2 Early Scots Data

The ESc data came from the following corpora: a) the *HCOS* and the *LAOS*²⁴ (all information relates to Version 1.1 of this corpus, unless stated otherwise) were used to select Scots samples totalling 497,572 words, the former corpus containing prose and verse texts, whereas the latter prose texts only; and b) the *TEAMS* was used to complement the data extracted from the *HCOS* and the *LAOS* with two ESc verse texts of over 1,800 words each (the description of the *TEAMS* will not be repeated here – see Section 2.2.1 for further information). The word counts of Scots texts in the corpora used for the ESc data set are given in Table 2 and have a total of 501,279 words.²⁵ Although there are other computerised Scots corpora available to carry out a language investigation, they were not selected for the purpose of this study for the following reasons: a) *Records of the Parliament of Scotland to 1707* by the University of St. Andrews contained a searchable database of the proceedings of the Scottish Parliament, but a number of such records were already included in the *HCOS* and the *LAOS*. Therefore, these additional records were not included in the ESc data set to avoid adding even more prose texts to already prose-dominated data sets; b) *Aberdeen Burgh Registers 1398-1511* by ScotlandsPlaces provided access to digital copies of Aberdeen’s Council Registers, yet a selection of these records could be found in the *HCOS* and the *LAOS*. Like in the preceding case, the burgh records are also written in prose; c) *The Breadalbane Collection* by the University of Edinburgh contains a collection of 16th century documents relating to the Campbells of Glenorchy, but were written in prose, too, and were also outwith the period of

²⁴ The total number of words of the texts in the *LAOS* is 401,672 (Williamson, personal communication, 11 August 2010). As the *LAOS* has no facility to provide word counts, I obtained these manually by setting up a *LAOS* task to search tagged texts by ‘Time & Space’ with the following parameters: TIME: 1375/1380-1419 and 1420-1500; SPACE: ALL counties. This returned a list with the names of dictionary tags that were matched up against word counts for each corpus file (Williamson, personal communication, 11 August 2010). I used the final sum from this list as a word count value for the *LAOS* data used in this study. However, during the counting process, I found no word count references to two dictionary files: [ABD/drumch01.tag](#) and [LAWS/butemsleges.tag](#). Nevertheless, details of the former are not provided with word counts for each corpus file (Williamson, personal communication, 11 August 2010), whereas upon attempting to access the latter one, the following error was returned: “We are sorry the page you requested was not found: [www.lel.ed.ac.uk/ihd/LAOS1/tagged_data/LAWS/butemsleges.tag](#)”.

²⁵ The lists containing ESc texts from the *HCOS* and the *TEAMS* used to carry out this study, together with the type of text, where this information is available, are provided in Table 61 and Table 62, respectively, in Appendix One. No list of texts and their respective word counts from the *LAOS* is available in the public domain, although Williamson (personal communication, 28 April 2014) has provided a file with tag count totals.

investigation covered in this study; and d) *The Corpus of Scottish Correspondence* contained Scots texts, but from the MSc period, namely 1500-1715.

Table 2

Word Counts of Scots Texts Used for the ESc Data Set per Period I and Period II

	<i>HCOS</i>	<i>LAOS</i>	<i>TEAMS</i>
Period I	1,396	30,291	3,707
Period II	94,504	371,381	-

As with every corpus, there are some significant differences between the *HCOS* and the *LAOS* in terms of the organisation of the samples. The former is an addition to the diachronic part of the *HCET* and provides samples enabling the researcher to study the last stages of the divergence within the Northumbrian *Sprachgebiet* (cf. Section 1.3.1), resulting in the rise of a distinctive Scots variety of ‘English’ (Meurman-Solin 1995: 50). Being part of the *HCET*, the *HCOS* also uses the COCOA header parameters to code various texts in respect to the date of the original, date of the MS etc., but with the following exceptions: a) the *HCOS* does not use the parameter denoting the dialect of the samples, with all texts being marked as written in Scots (this, however, is in contrast with the *LAOS*, where it is possible to search files according to different counties of Scotland, where this type of information is available. Nevertheless, there are also a number of sources in the *LAOS* which are referred to as ‘State Unlocalised’, ‘Diplomatic Unlocalised’ and ‘Other Unlocalised’); and b) The *HCOS* has “MEDIUM MS” under the parameter denoting the date of the MS in the COCOA header, perhaps suggesting that the source texts for the *HCOS* were not printed. Given the lack of specific information in relation to the date of the MS of every sample, I used the date of original of a text when grouping samples into Period I and Period II. Although this kind of approach to organising my data from this particular corpus was far from being ideal, given that it would result in unexpected and conflicting results, I would argue that it still presented a choice of dating by reference to the date in which the text was originally

composed, if the reference to the MS from which a samples came was either unknown or unavailable.

Unlike any of the above-mentioned corpora, the *LAOS* takes the form of an interactive website to search and display linguistic data from a corpus of 1,172 Scots texts dated 1375/1380-1500 (Williamson, personal communication, 28 April 2014). Every text in the *LAOS* is transcribed directly from an unedited text, MS or a facsimile (Williamson, personal communication, 29 Sept 2013). There are only prose documents and record texts in the *LAOS* (Williamson, personal communication, 23 September 2013), and the corpus refers to “primarily ‘local documents’, i.e. writings which can be localized from internal, non-linguistic references”.²⁶ Since this corpus operates differently, the organisation of headers in the *LAOS* is also different from those of the above-mentioned corpora: each text contains information about the source MS and a date code, if available (Williamson 2008: Tagged Texts: the headers).²⁷ On the basis of this, I grouped texts into Period I and Period II, covering the time spans of 1375-1420 and 1420-1500, respectively. This allowed a more reliable and thorough analysis of dialectal change, with differences between items under investigation, if any, being objectively measured. Finally, every text in the *LAOS* is lexico-grammatically tagged with “a description of the word or morpheme that seeks to capture its semantic and grammatical properties in the form of a ‘lexico-grammatical’ tag” (Williamson 2008: What is a Tagged Text). Having such lexico-grammatical tags had its advantage over the other computerised corpora, as every tag contained information about the semantic meaning, grammatical category and/or grammatical function of a word or morpheme.²⁸

²⁶ The list of all texts incorporated into the *LAOS* is not possible to provide here due to the spatial constraints of this study – see the *LAOS* website (Auxiliary Data Sets > Index of Sources; Bibliography) for an index of sources and references used in the compilation of the *LAOS*.

²⁷ Williamson (2008: Tagged Texts: the header) points out that in the *LAOS*, the New Year is recorded as observed on 25 March. Given this, any text in the *LAOS* dated between 1 January and 24 March is subject to an adjustment to a year beginning on 1 January. Thus, there are two forms of the date used in the compilation of the *LAOS*: one that makes this adjustment, expressed as a seven-digit code (which users of the website do not see; it is for processing) and one (which website users do see) that gives the date as 11 Feb 1466/67 (Williamson, personal communication, 19 August 2013).

²⁸ The *LAOS* contains editorial marks; these are used, for example, to account for instances where it may be impossible to obtain a reading from a piece of text or where characters cannot be deciphered or where the missing letters cannot be safely and successfully inferred (Williamson 2008: Tagged Texts: format of the text and tags). I have removed such editorial marks from the examples used in this study.

2.2.3 Representativeness

It is clear from the above sections that the corpora used for the data sets in this study can be described as not being fully representative for various reasons. Although there are a number of definitions of a corpus (Francis 1992: 17, Atkins, Clear & Ostler 1992: 1, Leech 1992: 116, Sinclair 1995), McEnery, Xiao & Tono (2006: 5; emphasis in the original) argue that “there is an increasing consensus that a corpus is a collection of (1) *machine-readable* (2) *authentic* texts (including transcript of spoken data) which is (3) *sampled* to be (4) *representative* of a particular language or language variety”. However, McEnery, Xiao & Tono (2006: 5) continue that although all scholars agree upon the first two qualities, the issue of what can be counted as representative attracts differing opinions. Both Francis (1979) and Sinclair (1996) also stress the importance of representativeness: the former defines a corpus as “a collection of texts assumed to be representative of a given language, dialect, or other subset of the language, to be used for linguistic analysis” (Francis 1979: 110), whereas the latter argues, “[a] corpus is a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language” (Sinclair 1996: Definitions: Corpus and computer corpus). Therefore, even though it may be tempting to define a corpus in linguistic terms, non-linguistic aspects or parameters associated with a corpus should also be taken into account. Below, I discuss matters associated with representativeness in relation to the size, type and genre of the data sets used in this study.

2.2.3.1 *Unequal Size of the Material*

The first issue concerned here is an unequal size of the material used in the NME and ESc data sets. But this study is not the only one faced with the problem of corpus representation, as issues linked to the representativeness of material (the size, type and availability of textual evidence) in general might be seen present in other studies, too. For example, Taylor (1917) concentrates on the use of *can* and *couth* in the works by Barbour and Blind Harry. Smyser (1967) and Richardson (1991a; 1991b) use works by Chaucer (the latter source also uses *Sir Gawain and the Green Knight*),

whereas Koziol (1932: 132-133) and Tajima (1975) base their analyses on poems by the *Gawain*-poet. But, given that **gan** and **can** have a discourse function in verse (Richardson 1991a: 76-89, Brinton 1996: 67-83), an investigation into the role of this verb and its variant in a limited amount of textual evidence may actually be seen as advantageous: not only is the size of a corpus used for any analysis dependent on the purpose for which it is intended, but also discourse functions of a linguistic item or linguistic items are elusive and may be highly dependent on text type, authorial preference etc. For example, Brinton (1996: 75) investigates a discourse function of **gan** in *Troilus and Criseyde* only, as “Chaucer uses the form more frequently than in any other text”. However, in contrast to the above-mentioned studies, other scholars extend their analyses to a wider selection of texts. This has its advantages, too – not only do such studies offer a comparison of the behaviour of **gan** and **can** in different texts, but also provide information on the frequency of occurrence of this verb and its variant which can be used to make generalisations about this verb and its variant in a variety and text, but also in reference to their function. For example, Terasawa (1974) looks into the origin, development and use of **gan** and **can**, as well as **beginnen** in a range of ME texts, giving a comprehensive overview of the frequency of these verbs with respect to date, dialect and genre.

However, the problem of an unequal size of the material from the corpora used in this study is more specifically related to the limited availability of textual evidence from before NME and the lack of this kind of evidence that would allow a lexicogrammatical analysis from before ESc, rather than my deliberate choice to investigate the verb **gan** and its variant **can**, as well as the verb **beginnen** in a smaller selection of texts. Nevertheless, Leech (1991) argues that the size of corpora used for analysis is not that important. For example, with a corpus, for example, of approximately 960,000 words long, Biber (1988: §4.3) appears to place emphasis on texts representing “a broad range of the possible situational, social, and communicative task variation occurring in the language”, rather than their length; Givón (1995: §1.8.2), on the other hand, uses small samples for an investigation of grammatical items.

This limited evidence, therefore, would have an impact on the balance of the material in the respective data sets and periods in this investigation, but also in relation to the types of text in the data sets under investigation. For example, although King (1997: 157) argues that “[f]rom the Early Scots period on, the range, quantity and subject matter of extant texts in Scots is impressively large”, copious amounts of textual evidence are unlikely to be immediately available at the start of the period that conventionally begins the linguistic history of ESc, i.e. post the Pre-literary Scots period (cf. Section 1.3.2), except for some place-names, personal names and some legal terms (Craigie 1931, Scott 2004: 214).

In order to overcome potential problems associated with the imbalance of textual material in the data sets and periods, I applied in this study a way of normalising frequency counts, with tables showing observed relative (normalised) frequencies, where applicable, in addition to observed absolute (raw) frequencies of occurrence of *gan* and *can*, as well as *beginnen* in the NME and ESc data sets, and the relative proportion of these verbs to each other (per-cent). However, higher relative frequencies of occurrence of these verbs in some texts, especially narrative ones (Terasawa 1974: 94-96, Brinton 1996: 69), does not necessarily mean that these verbs are frequently found in the data sets and periods, since those types of frequencies are dependent on the size of the material compared. For this reason, I provided the observed relative frequencies per 100,000 words in order to objectively compare the frequencies of occurrence of the above verbs across the data sets, periods, texts, samples etc. I will come back to this point in Section 2.4.1.

2.2.3.2 *Prose vs. Verse*

Gan and *can* frequently occur in verse texts (Terasawa 1974: 94-96, Brinton 1996: 69), although isolated instances of the former have been reported to occur, sporadically, in ME prose, too (Smyser 1967: 69, Brinton 1981: 164; 1988: 120; 1996: 69). With an exception of specialised corpora, a corpus is usually balanced in terms of prose and verse texts. However, the meaning of the term ‘balanced’ is relative, especially in relation to historical corpora where the types of texts and their genre will typically depend on what is available. An issue with representativeness

may, therefore, concern the skewed distribution of one type of text or the other in the material used in this study. As the notion ‘type of text’ can refer to the division of literary material into prose and verse, but also into different types of genre, I firstly deal with the former. In the *ICoMEP*, *HCOS* and *LAOS*, all texts are classified as prose by the compilers of these corpora,²⁹ whereas the *HCET* and the *TEAMS* have extracts of verse texts suitable for this study in various proportions. For example, in the case of the *HCET*, 90% of all samples classed as northern belong to the category ‘verse’ in Period I, with only *The Benedictine Rule* being classed as prose; in contrast, only 24% of all samples categorised as northern in this corpus are written in verse in Period II, namely *The Wakefield Pageants in the Towneley Cycle* and *The York Plays*. Given the lack of balance into prose and verse texts, and its potential implications in relation to the distribution of *gan* and *can* in general, I added northern and Scots verse texts from the *TEAMS*, as already mentioned in Sections 2.2.1 and 2.2.2: a) *Octavian* dated 1430-1440 (Thornton MS) and *Sir Eglamour of Artois* from ca. 1450 (MS Cotton Caligula A2), both written in a dialect of NME; and b) *The Life of St. Julian Hospitaller* and *St. Andrew and Three Questions* from *The Scottish Legendary* from ca. 1400 (Cambridge University Library MS Gg.ii.6), both characterised by a dialect of ESc (the word counts of the texts used in this study are given in Appendix One). The word counts of verse texts used in this study are given in Table 3.³⁰

²⁹ There can be occasional alliterative and formulaic tradition in prose texts, too; Markus (2006: 9) notes this in reference to texts contained in the *ICoMEP*,

[t]he texts are, therefore, relatively free from poetic stylisation (in spite of the occasional role of the alliterative and formulaic tradition). The language of this prose can be assumed to be closer than that of poetry (if not close) to the way language was really used in speech.

³⁰ This includes the word count of 372 words from *Meditatio de Passione Cristi* in *Richard Rolle of Hampole, Yorkshire Writers*. The occurrence of this verse extract in this prose text will be addressed in Chapter Six.

Table 3*Word Count of Verse Texts in the NME and ESc Data Sets per Period I and Period II*

	NME Data Set	ESc Data Set
Period I	24,487	3,707
Period II	29,862	-

2.2.3.3 Genre Type of the Material

Related to the type of text and its classification into either prose or verse is genre type, although genres do not rely on the modes (verse and prose) in which they are written (cf. also Tonkin 1992: 50-65, Turco 2000: 4 and Hiatt 2007 amongst others). Genre types are complex entities with a number of changeable characteristics (Whetter 2008: Redefining Medieval Romance, cf. Burrow 2008: 59-60). However, even though genre types have different characteristics, their purpose is the same as of any other discourse type. Burrow (2008: 60) argues, “[genre types] establish for readers, more or less precisely, what kinds of meaning they may expect to find in a text.” In this respect, rather than talk about abstract categories into which different texts should fit, we should view genre types as historical entities, whereby sets of conventions that characterise them at any one time alter from time to time. Therefore, every historical time has its own system of genres in which some kinds are dominant. In the case of the ME period, it is mainly the medieval romance³¹, also known as narrative poetry, beside the other genres also practised at the time, though to a lesser extent, namely the popular drama (the mystery play and the morality play) and a body of work commonly known as the “Middle English Lyric”. Burrow (2008: 70) argues,

[t]he chief strength of English literature in [ME] period, as in Old English, lies in narrative, rather than in lyric or dramatic writing. This is indeed the period within which English narrative verse reaches its apogee, in the Ricardian age, with the work of the Gawain-poet, Gower, and Chaucer; and

³¹ Dalrymple (2008: 55) maintains that the meaning of the term ‘romance’ is characterised by fluidity; originally, it referred to “any metrical narrative in the vernacular”, but by later Middle Ages, it came to “indicate more specific tales of adventure, magic, love and chivalry.”

throughout the period narrative proliferates in a rich variety of forms, prose and verse.

However, the term ‘narrative’ refers to a broad category of writings, defined in modern times as fiction, history, biography etc., but Burrow (2008: 61) proposes to apply the formal criteria of scope and scale. In relation to the former, which concerns the amount and complexity of the material, he argues that there are three degrees of it: a) histories (for example, Lazamon’s *Brut*, Malory’s *Morte Darthur*, *The Destruction of Troy* and *Cursor Mundi*), characterised by a division into episodes, offering a thematic unity, with a number of different characters, though not fully achieving the status of protagonist or hero; b) lives (for example, *Vita Sancti*, *Guy of Warwick*, *Lives of the Saints*), which like histories, deal with episodes, although loosely connected, with a single protagonist, who ensures a rudimentary unity; and c) tales (for example, *Gest of Robyn Hode*), which often consisted of histories and lives combined together, with the tale focussing on an excerpt from one’s life, such as a miracle or a love story, whereas the lives describe one’s complete life, for example the life of a saint (Burrow 2008: 60-64). In relation to scale, which deals with “the degree of detail with which [the] material is presented”, Burrow (2008: 65) argues that narratives will differ in length, from part to part. Considering this, he speculates that there should be more opportunities for a detailed narrative in tales than in histories or lives, given that there is less of a story to tell.

While the corpora used in this study have a variety of genre types, some genres do indeed tend to ‘dominate’ either or both of the two varieties, or the periods. However, not only does this dominance appear to be linked to the type of material produced at the time, but also to the availability of extant material today. For example, in the *HCOS*, six texts in Period II (35%; 51,305 words) out of all ten (65%; 95,900 words) in Period I and Period II are classed as ‘Education Treatise – Instruction’, with their sub-genre being either ‘religious’ or ‘secular’. Furthermore, even though some genres may seem to dominate in a particular corpus and/or period of the corpora used in this study, there may also be problems with a consistent classification of texts into genre types. For example, while the *HCET*, the *ICoMEP* and the *HCOS* share genre labels, the *LAOS* uses a different system of labelling

documents, such as ‘book/record/court’ to describe *Ayr Burgh Court Book* or ‘notarial protocol book’ to refer to *Protocol Book of Sir Thomas Crawford*. Finally, literary texts can consist of a mixture of prose and verse, making such evidence more complicated to classify into their respective categories. For example, *Meditatio de Passione Cristi* with two instances of **gan** is part of Rolle’s prose work *Richard Rolle of Hampole, Yorkshire Writers* in the *ICoMEP* but contains verse-like characteristics (metre, internal and end rhymes), even though it is written as prose (cf. Appendix Two). Therefore, in order to remain consistent and to avoid misclassification, overlapping and bias, I use the term ‘genre’ to refer to a classification of texts based on external criteria adopted by the compilers of the corpora; I also maintain genre labels used in the *HCET*, the *ICoMEP*, the *HCOS* and the *LAOS*, as outlined in Appendix One.

To make the matter more complicated, no system of formal classification into different types of texts is offered by the editors of *TEAMS*, although the introductions provided for each text offer detailed descriptions of the background of a text, including some clues as to the types of genre of these texts. On the basis of this, the texts used in this study were assigned the following classification: ‘romance’ for *Sir Eglamour of Artois* and *Octavian*, given frequent references to this genre type in Hudson’s (2006a; 2006b) introduction to both texts in the *TEAMS*, and ‘legend’ for *St. Andrew and the Three Questions* and *The Life of St. Julian Hospitaller*, following references to this genre type in Whatley, Thomson & Upchurch’s (2004a; 2004b) introductory remarks on both works. Even though the latter two sources are classified in this study under the category of ‘legend’, there is a similarity between ‘legend’ (of the saints) and ‘romance’ in that the former can be seen as containing romance narrative. However, as the same could be said about ‘historical account’ (as represented in this study by *Cursor Mundi*) and ‘romance’, and ‘homily’ (as illustrated by *The Northern Homily Cycle*) in that such texts can be part of larger texts or contain extracts belonging to a different category altogether, I decided to treat ‘legend’ here as a separate category, together with the division into ‘history’ and ‘homily’. As is the case with the classification assigned by the authors or editors

of the *HCET*, the *ICoMEP*, the *HCOS* and the *LAOS*, these types of labels are also used in this study and provided in Appendix One.

2.2.4 Spelling Variation

When using corpora, especially electronic ones, to conduct a language study, there is a note of caution about their use. The irregular nature of spelling in historical texts is an integral part of the diversity of sources in any analysis. Before annotation of the data sets, I therefore compiled a list of forms of the verb *gan* and its variant *can*, as well as the verb *beginnen* (including inflected forms, where different and/or available) in NME and ESc. I did that using reference sources, namely the *DSL* (s.vv. *begin* v.; *can* v. 2; *gan* p.t.), the *MED* (s.vv. *biginnen* v.; *can* v.; *ginnen* v.) and the *OED* (s.vv. *begin* v.¹; *can* v.²; *gin* v.¹), since they all provide a comprehensive list of such spelling variants. However, I encountered two problems with this kind of approach: a) there would be a risk of missing and/or misclassifying some forms, for example, where the quality of edited versions was in question because of omission or misspelling; and b) the list of spelling variants available in the above reference sources may be incomplete, especially in relation to NME and ESc as two relatively under-researched varieties. To overcome these obstacles, I manually scanned line by line the northern samples from the *HCET*, the *ICoMEP* and the *TEAMS*, and Scots samples from the *HCOS* and the *TEAMS* for the potential presence of *gan* and *can*, as well as *beginnen*. While doing so, I was able to check any dubious instances of these verbs in order to either confirm or reject their suitability for analysis. However, this task proved very laborious and time-consuming, but, ultimately, minimised the risk of potential omission and misclassification of the above verbs if only searching the corpora electronically.

It is inevitable that some items may be homographs with *gan* and *can*, as well as *beginnen* in the NME and ESc data sets. Fortunately, with the availability and help of the above-mentioned reference sources, homographs can be discriminated between on the basis of meaning of each form. For example, forms like <gonne> being the second person singular of *gan* could potentially be confused with <gonne> of the verb *gōn* ‘to walk; also, be able to walk’, ‘to travel (by walking, riding, sailing,

etc.), proceed, go, move’ and ‘to go to (sb.); travel over (a place), traverse; sail to or along (a coast)’ (*MED* s.v. *gōn* v. 1a-3). In a similar way, *can* as a variant of *gan* could be miscategorised as a modal (auxiliary) verb, given the fact that their forms are similar in some dialects and identical in others. For example, <con(ne)> being the first and third person singular form of *can* meaning ‘a modal verb stressing the fact of an act or event: (a) did (do sth.); also, began (to do sth.)’ can have the same orthographic form as <con> ‘to have ability, capability, or skill: be able (to do sth.), be capable (of doing sth.), know how (to do sth.)’ (*MED* s.v. *can* v.; *cōnnen* v. 1). This potential confusion is also pointed out by the *OED* (s.v. *can* v.²), which maintains about the variant *can*,

[t]he word was probably often perceived as a special use of [modal (auxiliary) verb *can*] In Older Scots, forms of the past tense of [the modal (auxiliary) verb] like *couth*, *cowd*, *could* ... frequently replaced the original [*can* as a variant of *gan*] in its function as a periphrastic auxiliary of the past tense, in order to remove an apparent irregularity where what was perceived as a present stem of [modal (auxiliary) verb] was used as a past tense in the periphrastic construction.

Finally, the verb *beginnen* with the ingressive meaning could be confused with *beginnen* meaning ‘to entrap, ensnare’ (*OED* s.v. *begin* v.²), as in (1) below. I, therefore, made an attempt to establish the meaning of such doubtful instances by using reference sources like the *DSL*, the *MED* and the *OED*.

- (1) “*Laban*,” *sco* *said*, “*allas for sinn*
 “*Laban*,” *she* *said*, “*alas for sin*
 Qua wend *he wald* *þus me* *biginn*.”
 who thought *he would* *thus me* *entrap*.”
 (1350-1420 *Cursor Mundi* p. 229; *HCET*)

““*Laban*,” *she* *said*, “*who thought he would entrap me committing sin in this way*.””

2.3 Classification of Items under Investigation

The reliability of the results, and, ultimately, the analysis, very much depend on the accurate and systematic classification of the items to be investigated (in addition to relying on the scholarly quality of the editions of texts providing basis for the samples included in the corpora). Therefore, categorising items incorrectly in this

investigation could have serious repercussions on the outcomes. In what follows, I outline semantic and syntactic criteria by which I have classified the items investigated in this study.

2.3.1 Semantic Criteria

In most cases, *gan* and *can* have a non-ingressive meaning, as in (2) below and (3) below (*DSL* s.vv. *can* v. 2; *gan* p.t., *MED* s.vv. *can* v., *ginnen* v. 3b, *OED* s.vv. *can* v.² I 1a; *gin* v.¹ 1a, Taylor 1917, Funke 1922: 14, Koziol 1932: 133, Mustanoja 1960: 611, Kerkhof 1966: 156-158, Visser 1969: 1379, 1571-1575, 1577-1579, Brinton 1981: 159, 161; 1983: 235, 236; 1988: 120, Denison 1993: 322, Ogura 1998: 300). Instances of this verb and its variant occurring with an ingressive meaning have also been identified, as in (4) below, although less frequently so than those with non-ingressive meaning (*DSL* s.vv. *can* v. 2; *gan* p.t., *MED* s.vv. *can* v.; *ginnen* v. 1-3a, *OED* s.vv. *can* v.² I 2; *gin* v.¹ 1a, Einkenkel 1891: 89-90, Taylor 1917: 576, 583, 589, Beschorner 1920: 12, Häusermann 1930: 18, Koziol 1932: 131-133, Mustanoja 1960: 613, Kerkhof 1966: 154-156, Visser 1969: 1379-1380, 1573, Brinton 1981: 159; 1983: 235; 1988: 120). However, despite a clear-cut division into either non-ingressive or ingressive, Visser (1969: 1379) points out that the distinction between the meanings of the above-mentioned verb and its variant is “largely arbitrary” (cf. Einkenkel 1891, Taylor 1917: 55, Kerkhof 1966: 158-159, Brinton 1981: 159, Denison 1993: 322). Overall, the ingressive meaning in some instances of *gan* and *can* is well documented, but it is also clear that in many cases, this ingressive meaning is bleached. The meaning of this verb and its variant sometimes depends on fine points of interpretation. In what follows, I, therefore, examine the criteria used to distinguish non-ingressive uses of *gan* and *can* from ingressive ones.

- (2) *To bed thai gun go*
 To bed they did go
 (c1330 *Amis and Amiloun* 1161; *OED* s.v. *gin* v.¹ 1a)

‘They went to bed’

- (3) A new liueland gan [Cotton Vespasian A.3 MS *cun*;
Oxford, Bodleian Library, Fairfax MS 14 *con*; Trinity College,
Cambridge MS R.3.8 *dud*]³²
A new living did
he *bigin*.
he begin.
(a1400 *Cursor Mundi* 2009 (Göttingen MS Theol. 107); *MED* s.v.
ginnen v. 3b)

‘He began a new life.’

- (4) And er we weoren war, with Wit *conne* we
And before we were aware, with Wit began we
meetten
meet
(n.d. PPl. A [I] [Vrn] 9.109; Brinton 1988: 157; translation in the
original)

‘And before we were aware, we began to meet with Wit’

A number of attempts have been made in order to differentiate between non-ingressive and ingressive meanings of *gan* and *can* in ME and Scots, with various degrees of success and acceptability. One way of highlighting the occurrence of non-ingressive instances of this verb and its variant has been by comparing the meaning of *gan* and *can* to that of periphrastic *do* found in non-emphatic, affirmative and declarative clauses. For example, Mustanoja (1960: 602-603) argues that the use of periphrastic *do* is in a “colourless and unemphatic” way, making it similar to “that of [*gan*] in the periphrastic preterite”. Visser (1969: 1572), on the other hand, demonstrates that the use of *gan* is interchangeable with periphrastic *do*, as in (5) below, whereby the meaning of both verbs appears to be causative. However, Brinton (1983: 237) maintains that “clearly these periphrases have different meanings [in (5) below]; the *do* periphrase is causative, and the *gannen* [sic] periphrase is ingressive.” Further support in favour of the view that these verbs are interchangeable comes from the fact that they alternate in different versions of MSS

³² The *eLALME* provides the following information regarding the dialect and dating of these MSS: The Göttingen MS Theol. 107 – “forms of probably SE Lincs origin (Crowland area)”, but “from ca. lines 11000 on, the language is thoroughly and consistently northern (probably YWR [Yorkshire, West Riding])” (first quarter of the 14th century); Cotton Vespasian A.3 MS – “YWR [Yorkshire, West Riding]” (probably the 14th century) (LP18); Oxford, Bodleian Library, Fairfax MS 14 – “LAN [Lancashire]” (ca. 1400) (LP6); and Trinity College, Cambridge MS R.3.8 – “STF [Staffordshire]” (the 14th century) (LP36).

of *Cursor Mundi*, as in (3) above (*MED* s.v. *ginnen* v. 3b, cf. Visser 1969: 1572), although Mustanoja (1960: 614) quotes Deutschbein (1917: 6), who finds that in the Cotton MS of the above work, periphrastic *do* is replaced by the simple preterite (in later MSS of this work, by the verb *gan*), whereas Zilling (1918) reports an opposite tendency.

- (5)
- | | | | | | | | |
|-------------|--------------|-------------|---------------|------------------|------------|--------------|---------------|
| <i>For</i> | <i>with</i> | <i>that</i> | <i>oon</i> | <i>encresede</i> | <i>ay</i> | <i>my</i> | <i>fere,</i> |
| For | with | that | one | increased | always | my | fear, |
| <i>And</i> | <i>with</i> | <i>that</i> | <i>other</i> | <i>gan</i> | <i>myn</i> | <i>herte</i> | <i>bolde,</i> |
| And | with | that | other | began | my | heart | embolden, |
| <i>That</i> | <i>oon</i> | <i>me</i> | <i>hette,</i> | | | | |
| That | one | me | hot, | | | | |
| <i>That</i> | <i>other</i> | <i>dide</i> | <i>me</i> | <i>colde</i> | | | |
| That | other | did | me | cold | | | |
- (Chaucer *Parliament of Foules* 143; Brinton 1983: 237; translation in the original)

‘for with that one my fear increased, and with that other began my heart to be bold; that one made me hot, that other made me cold’

In addition to similarities in meaning between the verb *gan* and its variant *can*, as well as periphrastic *do*, a number of sources identify parallel functions performed by these verbs in ME. For example, Mustanoja (1960: 602-603) sees the role of both *gan* and *do* primarily to enable the verb in infinitive to provide rhyme at the end of a line, as in (5) above and (6) below, respectively (cf. also Ellegård 1953: 146, 208, Smyser 1967: 82-82, Visser 1969: 1498, Kroch 1989: 141 and Rissanen 1991: 333 amongst others). In these examples, the verbs *bolde* and *sprede* in infinitive occur in rhyme end position, although it seems that the meaning of *do* is causative, rather than periphrastic, in both examples. However, Mustanoja (1960) and Visser (1969) also argue for a secondary role of the verb *gan* and *can*, as well as periphrastic *do* in ME verse, although both sources differ in the type of function these verbs perform. For example, Mustanoja (1960: 602) maintains that the role of all of these verbs is to emphasise the event or situation denoted by the verb in infinitive, Visser (1969), on the other hand, argues that they make an extra syllable available to fill a line of metre with the required number of syllables. Given that the similarities in meaning and function between these verbs, as well as the fact that the rise in the use of periphrastic *do* coincides with the decline in the application of *gan* and *can* in the

15th century, a number of sources have concluded that the verb *gan* and its variant *can* must be the precursors of periphrastic *do* (Taylor 1917: 591, Funke 1922: 26, Visser 1969: 1572-1573, cf. Brinton 1983: 237, Sims 2014: 67).

(6) *whan Phebus doth his bryghte bemes sprede*
 when Phebus does his bright beams spread
 (Chaucer *Troilus and Criseyde* 2.54; Mustanoja 1960: 602)

‘when Phebus makes his bright beams radiate’

However, although there is some degree of overlap in meaning and function of *gan* and *can*, as well as *do*, whether the former two are the precursors of the latter depends on “one’s understanding of the origin, functioning, and development of *do*” (Brinton 1983: 237). I start off with the origins of *gan* and *can*, which derive from OE *onginnan* and OE and eME *aginnan* (*MED* s.vv. *aginnen* v.; *ginnen* v., *OED* s.vv. *gin* v.¹; *ongin* v., Visser 1969: 1372). The *OED* (s.v. *begin* v.¹) argues that the Germanic root meaning of **ginnan* was ‘to cut open, open up, begin, undertake’, cognate with the OE *ginnan* ‘to gape, to yawn’. Inherent in this root meaning was a sense of motion and, according to this source, the transition of meaning from ‘open up’ to ‘begin’ was a frequent one. Therefore, there were the following compounds of **ginnan* in Germanic languages: Gothic *du-ginnan* ‘to begin’, Old High German *in-ginnan* and Middle High German *en-ginnen*, beside OE *onginnan* and OE and eME *aginnan*, ‘to begin’ (Bosworth 2010 s.vv. *aginnan* v.; *on-ginnan* v. I, *OED* s.vv. *agin* v.; *ongin* v., cf. Brinton 1981: 89-108, 145-148; 1988: 117, Sims 2014: 60).

Although instances of *onginnan* and *aginnan* were recorded with ingressive sense in OE, both verbs also showed some bleaching of meaning in this variety already³³ (Wuth 1915: 56-57, Funke 1922: 5, Mustanoja 1960: 610-611, Visser 1969: 1372, 1580, Brinton 1988: 160-161, Los 2000; 2005: 88-99, cf. Sims 2014: 66-67). For example, Los (2000: 261) argues that the bleaching of meaning of OE *onginnan* is linked to the type of infinitive with which this verb is recorded in Ælfric, “[b]are infinitives after *onginnan/beginnan* no longer allow temporal segmentation in Ælfric,

³³ OE *onginnan* is also found with other senses, which, nevertheless, appear to be lost during the course of OE: ‘to attempt, endeavour (with infinitive)’, ‘to act strenuously’ and ‘to make an attempt upon, to attack’ (Bosworth 2010 s.v. *on-ginnan* v. II-IV).

and its selection signals to the hearer that the action described in its complement should be viewed as being completed without interruption”.

An alternative view of the origins of *gan* and *can* in ME is proposed by Homann (1954). While it is generally accepted that *gan* and *can* derive from OE *onginnan* and OE and eME *aginnan*, Homann (1954: 397-398) ascribes the use of *gan* to a similar construction in French. She argues that Chaucer and his contemporaries worked with what she terms as “a transitional language” and used *gan* with infinitive on the imitation of French reflexive constructions “s’en aller” beside “aller”,

[a]s part of the transition, the earlier ability to form causative verbs becoming weakened, compound tenses were formed. At the same time something of a causative sense was gained by the development of ‘did,’ ‘let,’ ‘make,’ and ‘gan’ as auxiliaries ... Often the English writer translates the French reflexive into his periphrastic ... The translator may have felt that the simple preterit did not convey the movement of the French reflexive or that he needed to indicate a greater immediacy and precision of tense for his narrative.

The above claim pointing to the French influence suggests that the occurrence of *gan* and *can* in ME and Scots is due to borrowing. The expectation would, therefore, be to see the use of this verb and its variant in higher numbers in works where the French influence is strong. However, both Beschorner (1920: 9-10) and Mustanoja (1960: 613) argue that Chaucer seldom uses *gan* in his earlier poems, such as *The Book of the Duchess*, written under the strong influence of French. Mustanoja (1960: 613) continues that it is not until Chaucer’s journey to Italy that *gan* appears more frequently in his works, for example in *The House of Fame*, *The Parliament of Fowls* and *The Legend of Good Women*, reaching its peak in *Troilus and Criseyde*. Therefore, by my count, it is possible to identify 14 instances (observed relative frequency of 161 instances per 100,000 words) of *gan* with either the plain infinitive or the *to*-infinitive in *The Book of the Duchess* (8,672 words); on the other hand, there are 59 instances (observed relative frequency of 444 instances per 100,000 words) of this variant with the plain infinitive, the *to*-infinitive and the *for-to* infinitive in *The House of Fame* (13,302 words) (this constitutes an increase of 176%) (cf. Skeat 1915: 83-96, 326-348). Nevertheless, while Beschorner (1920) and Mustanoja (1960) demonstrate that the occurrence of *gan* is lower in French-influenced works, they provide no explanation why this is the case. I would,

therefore, argue that the reason for the lower frequency of occurrence of *gan* and *can* in such works is because of their genre. This verb and its variant are associated with narrative texts (Terasawa 1974: 94-96, Brinton 1996: 69), but medieval romance “had no learned Latin tradition behind it” (Burrow 2008: 63). It is, therefore, unlikely that *gan* and *can* in auxiliary uses arose under the French influence in ME.

In relation to the origins of periphrastic *do*, various hypotheses have been central to past investigations. Most of the earlier theories are summarised in Visser (1969: 1488-1571), Fischer (1992: 267-276) and Denison (1993: 264-291), and so they will only be briefly repeated here. While evaluating each one of them would take more space than this thesis allows, the point of listing various theories on the origins of this construction is to highlight a different provenance of periphrastic *do*. Most linguists agree that periphrastic *do* develops from the earlier uses of *do* (Fischer 1992: 269) and so it is clear that this verb is unrelated to *gan* and *can* in terms of its origins. Furthermore, the most relevant theories on the origins of this construction are those that attempt to account for the presence of this verb in affirmative, non-emphatic and declarative sentences. The development of periphrastic *do* has, therefore, been proposed to be out of: a) the use of *do* as a causative verb (Deutschbein 1917: 810, Zilling 1918: 8, Funke 1922: 1, Ellegård 1953: 15-27 amongst others), which has been put forward most frequently as a likely source for this construction type; b) a development of the lexical verb *do* (Koziol 1936: 460, Marchand 1938/39, Marchand 1939: 123 amongst others); and c) the cataphoric use of factitive *do* (Visser 1969: 1492, 1497): c) the corresponding use of *faire* in French (Earle 1887: §586) or to Celtic influence (Preusler 1956: 334-336), both of which views are no longer upheld (Ellegård 1953: 92, 119-120, Visser 1969: 1495-1496), even though Ellegård (1953) argues that the French construction might have had only an indirect influence on the development of periphrastic *do*. In addition to the above theories, Denison (1985; 1993: 279-281) argues for the syntactic origin of periphrastic *do*; Poussa (1990) proposes that periphrastic *do* develops out of language contact with OE and Celtic, whereas Tiekens-Boon van Ostade (1985; 1987; 1989; 1990) suggests that it arises out of language contact with Germanic.

Not only is the geographical distribution of the verb *gan* and its variant *can* different from that of periphrastic *do*, but also the timing of the occurrence of these verbs presents problems, highlighting the disparity between these verbs in terms of use. For example, Sims (2014: 67) maintains that *gan* appears in the north in 1200 and in the south in 1300. Einkenkel (1891: 89), Wuth (1915: 54-57), Funke (1922: 7), Beschorner (1920: 9-10), Häusermann (1930: 18-20), Mustanoja (1960: 610-615), Visser (1969: 1379-1380, 1571-1575), Brinton (1983: 235; 1988: 117-122; 1996: 67, 69, 71) and Denison (1993: 321-322) amongst others demonstrate that both *gan* and *can* are frequent throughout the ME period, with auxiliary uses of *onginnan* being already found in OE (Brinton 1988: 159-161, Warner 1993: 117, Los 2000; 2005: 88-99, Sims 2008: 121-122; 2014: 66-67). Periphrastic *do*, on the other hand, is recorded in declarative uses in the 13th century rhyming verse from the southwest or the west of England, and in negatives, interrogatives and in inversion in the 14th century (Mustanoja 1960: 603, Visser 1969: 1522-1541, Denison 1993: 264), with no certain examples of this verb being recorded in OE (*OED* s.v. *do* v. 32).³⁴ Visser (1969: 1498) also speculates that periphrastic *do* must have been in spoken use “in pre-Conquest times” already, although it is impossible to prove this usage was indeed in speech before the 11th century,

[t]hat it did not occur in the existing documents of the time is ascribable to the fact that with the construction without *do* (e.g. * ‘he weop’) everything that was to be expressed could be expressed and that no notion whatever was added by displacing it by the periphrastic pattern (e.g. * ‘he dyde wepan’). And it was this superfluity that withheld writers from using it. When, however, in the course of the thirteenth century non-alliterative verse came to be the vogue, poets began to discover ... the advantages of the periphrastic pattern offered in writing verse [, namely the function of periphrastic *do*, as outlined in the preceding paragraphs]

Mustanoja (1960: 603-604) adds that the earliest instances of periphrastic *do* in prose are from 1400, but this construction remains infrequent in the east than in the west until the end of the 15th century. He continues that that it is not found in 15th century prose in the north, remaining rare in this type of writing there in the 16th and the 17th

³⁴ Although isolated instances of *gan* occur in prose, too (Smyser 1967: 69, Terasawa 1974: 94-96, Brinton 1981: 164; 1988: 120; 1996: 69), Visser (1969: 1498) argues that “[i]t took some time for prose – which had nothing to gain by it – to begin to use the periphrastic pattern *do*”. However, Mustanoja (1960: 603-604) claims that periphrastic *do* is not found in prose written in the north in the 15th century, while still occurring relatively rare in this type of genre in the 16th and 17th centuries.

centuries. However, he argues that this construction becomes more popular in the south during the first half of the 16th century. Incidentally, this increase in frequency of periphrastic *do* is happening at the same time as a decline of the verb *gan* in the early part of the ModE period. Visser (1969: 1572-1573), for example, notes that by the 15th century, *gan* and *can* are characterised by a distinct decline in the frequency of occurrence, whereas in the 16th century, both are seen as obsolete (cf. Mustanoja 1960: 614). On the basis of this type of information, Denison (1993: 322) comments that this verb and its variant are “effectively” replaced by periphrastic *do*, suggesting at the same time that this is the reason for the disappearance of *gan* and *can*.

However, Visser (1969: 1573) concludes that it is not possible to speculate about a reason for the decline of this verb and its variant, even if there is a degree of overlap in the timing of these constructions; nevertheless, he does not seem to entirely rule out the potential influence of periphrastic *do* on the decline of the verb *gan* and its variant *can*, “[t]he cause of [*gan* and *can*] falling into disuse is unknown; it may, perhaps, be seen in the reluctance to apprehend *can* as a preterite, where *did* was available.”

Differences in the origins and development of the verb *gan* and its variant *can*, and periphrastic *do* may also be reflected in the use of these verbs in literary works. For example, Mustanoja (1960: 603) notes that “[t]he construction [involving periphrastic *do*] is by no means common in Chaucer, who seems to prefer periphrastic *gin*”, whereas Visser (1969: 1572) argues that “[t]hat periphrastic *did* and periphrastic *gan/can* were interchangeable appears from c1381 Chaucer, Parl. of Foules 143”, although *gan* is not present in Chaucer’s *The Tale of Sir Thopas* (Funke 1922: 26, Smyser 1967: 83, Brinton 1996: 71). Although the north is distinctly affected by the distribution of periphrastic *do*, Macafee (1992/93: §8.13), on the other hand, notes that in MSc, *do* is used in Chaucerian-influenced styles, whereas *gan*, *can* and *couth* in more native styles. Although she provides no further details, it is well known that English poetry and prose was present in Scotland from the 14th century or earlier, with Scottish poets of the 15th and the 16th centuries imitating Chaucer’s style, but also English spellings and locutions (P. Blank 1996: 154). Yet, it will be recalled from the previous paragraphs that “[periphrastic *do*] is by no means

common in Chaucer, who seems to prefer periphrastic *gin*] (Mustanoja 1960: 603). Therefore, it could be that the distribution of *do*, and *gan* and *can* depends on the nature of semantic and stylistic constraints in such works, with these verbs being used to convey separate meanings and being used in different functions. Overall, given the contrasting origins and divergent development of the verb *gan* and its variant *can*, and periphrastic *do*, evidence in support of the similarities in meaning and function between these verbs remains inconclusive. From the methodological point of view, this ambiguous evidence in support of *gan* and *can* being similar to *do* could, therefore, have an impact on selecting non-ingressive instances of *gan* and *can* in order to determine the extent to which this verb and its variant have undergone semantic bleaching.

While the above claims to describe the non-ingressive meaning of the verb *gan* and its variant *can* in an attempt to separate such instances from those where their meaning is ingressive are based on non-systematic observations, proposals made by Brinton (1981; 1983; 1988) to distinguish between non-ingressive and ingressive instances of *gan* and *can* are formulated within the framework of a theory of aspect. They can, therefore, be systematically applied to the above verb and its variant when attempting to identify their meaning in the data sets under investigation: Brinton (1981: 169; 1983: 238-239; 1988: 155), for example, maintains that an ingressive interpretation of *gan* and *can* is not possible when this verb and its variant occur with a verb in infinitive denoting a punctual event or situation, as in (7) below. In this example, the second instance of *gan* is used with a punctual (achievement) verb *mysse* ‘to notice the absence or loss of; to perceive that (a person or thing) is not in the expected or accustomed place’ (*OED* s.v. *miss* v.¹ III: 16, cf. Wuth 1915: 55, Beschorner 1920: 13-14, Funke 1922: 12, 24, Kerkhof 1966: 156-158).³⁵ In such cases, the event or situation happens in a moment without being portrayed as having a distinct beginning, middle or end. It is, therefore, not possible to focus on the start of the event or situation in question, as it coincides with its conclusion (Brinton

³⁵ The verb *awake* in this example may also be seen as a punctual (achievement) verb (resulting in the meaning of *gan* as non-ingressive) in that it focusses on the moment of being asleep. On the other hand, *awake* may be seen as an accomplishment verb in that it can involve duration, with the endpoint of coming out of sleep being approached incrementally.

1983: 238). On the other hand, an ingressive interpretation is possible when a verb expresses a durative event or situation, isolating the moment of inception from the rest of the event or situation described, as (8) below (Brinton 1983: 238-239; 1988: 155). However, I would argue that a non-ingressive interpretation of *gan* in this example would be better, since *tak[ing] the letter* involves a terminal point (I have kept here the translation provided in Brinton 1988).

- (7) *And gan awake, and wente hire out to*
 And did/began to awake, and went her[self] out to
pisse, And cam agayn, and gan hir cradel
 urinate, And came again, and did her cradle
mysse
 miss
 (Chaucer *Canterbury Tales The Reeve's Tale* A.4215-4216; Brinton 1988: 156)

‘And did/began to awake, and went out to urinate, and came back, and missed her bed’,³⁶

- (8) *And Pandarus gan him the letter take*
 And Pandarus began to him the letter take
 (Chaucer *Troilus and Criseyde* 2.1318; Brinton 1988: 155;
 translation in the original)

‘And Pandarus began to take him the letter’

In addition to the above criteria based on whether the verb in infinitive found with *gan* and *can* denotes a punctual event or situation, non-ingressive instances of *gan* and *can* with a verb in infinitive denoting this kind of event or situation can occur with certain types of subjects and objects. For example, Brinton (1981: 172) argues that “when the object or subject of the punctual verb is a plural count or mass noun, it seems easier to interpret [the verb *gan* in an ingressive sense]”, whereas in Brinton (1983: 240), she slightly refines this proposal, “[w]hen the object or subject of the construction with [*gan*] and a punctual verb is a plural count noun or a mass noun, however, rather than a single count noun, an aspectual [ingressive] interpretation of [*gan*] seems possible”. Examples provided by Brinton (1983: 240) include the

³⁶ Brinton (1988: 156) translates the above as ‘And began to awake (awoke), and went out to piss, and came (in) again, and missed (NOT began to miss) her cradle’; I have provided an alternative translation.

punctual (achievement) verbs *quelle* ‘kill’, as in (9) below, *mete* ‘meet’ and *deffie* ‘defy’ with *gan*, but also *sle* ‘slay’ with *ginnen* (present indicative), as in (10) below. The presence of plural subjects or plural objects means that the event or situation denoted by the verb in infinitive is presented as durative, i.e. an iterative series (1983: 238-239; 1988: 119-120), and so, reinterpreted as ingressive, even though the verb in infinitive is a punctual (achievement) verb. However, Brinton (1988: 157) claims that this diagnostic can be applied in “certain contexts”, whereas it is omitted altogether in Brinton (1996). I conclude that if the subject or object is a plural count noun or a mass noun, then it does not matter if the verb in infinitive is a punctual (achievement) verb or any other type of verb, as in such cases, an ingressive interpretation is likely. I also show in Chapter Five that only in some cases, it may be helpful to establish the non-ingressive meaning of *gan* and *can* in the NME and ESc data sets on the basis of this diagnostic; I, therefore, use this diagnostic with caution.

- (9) *Pat folc hi gunne quelle*
 That folk they began to kill
And churchen for to felle
 And churches for to fell
 (*King Horn* 61-62; Funke 1922: 13, as cited in Brinton 1981: 172)

‘They began to kill people and demolish the churches’³⁷

- (10) *He gynnyth to sle with dilfull dynt*
 He begins to slay with sorrowful blow
A thowsaund men or euer he stynt
 A thousand men before ever he ceases
 (a1475 *The Siege of Troy* (Harley 525) 1138; Brinton 1983: 240;
 translation in the original)

‘he begins to slay with cruel blow a thousand men before he stops’

Finally, other linguistic items in a sentence may prevent an ingressive interpretation of *gan* and *can*. For example, Brinton (1981: 173) proposes that the non-ingressive interpretation of this verb and its variant is likely when this verb and its variant occur with “durative (or iterative) adverbials [, or adverbs in Brinton (1983; 1988)] which are antithetical to a focus on the moment of inception”, such as *ful oft* ‘very often’ in

³⁷ Brinton (1981: 172) translates the above as ‘They began to kill people and to fell the churches’; I have provided an alternative translation.

(11) below (cf. Brinton 1983: 240).³⁸ Therefore, in line with these proposals, the presence of the iterative adverbial *ful oft* ‘very often’ in this example makes an interpretation of the variant *can* as non-ingressive, as an iterative adverbial is antithetical to a focus on the moment of inception (Brinton 1981: 173).

(11) *his cher ful oft con chaunge/*
 his mood full often did change /
at chapel er he myzt sene
 that chapel before he might see
 (*Sir Gawain and the Green Knight* 711-712; Mustanoja 1960: 611, as cited in Brinton 1981: 173)

‘his mood changed very often before he saw that chapel’,³⁹

Overall, therefore, using the semantic criteria outlined in this section, it can be said that instances of *gan* and *can* are non-ingressive, i.e. in auxiliary use, if this verb and its variant are found with a punctual (achievement) verb in infinitive or when they occur with a durative or iterative adverbial, unless the subject or object of this verb and its variant found with a punctual (achievement) verb in infinitive is a plural count noun or a mass noun, in which case an ingressive interpretation of *gan* and *can* is likely in some cases. However, Brinton (1981; 1983; 1988; 1996) primarily relies on semantic (or functional) criteria to separate the auxiliary uses of *gan* from the other, namely where this verb has a full, ingressive, meaning; this kind of criterion is also used in other sources, such as Mustanoja (1960), to distinguish between auxiliary and non-auxiliary uses of this verb. Nevertheless, sources like Mustanoja (1960) and Brinton (1981; 1983; 1988; 1996) do not take into account formal criteria, for example, as defined by Warner (1993: 103) in relation to auxiliary verbs, “[s]ubcategorization for the plain infinitive, not for the to-infinitive” is a criterion determining auxiliaryhood. Therefore, I would disagree with Brinton (1988: 111) when she implies throughout her study that auxiliation is primarily a semantic process, “[v]iewed as a semantic process, auxiliation raises two issues: one about the choice of verbs to become auxiliaries and the other about the semantic changes

³⁸ However, although the above sources use adverbials of duration or iteration as a criterion to distinguish between ingressive and non-ingressive meanings of *gan* and *can*, the use of adverbials of time-at-which can be applied to constructions conveying both ingressive and non-ingressive meaning.

³⁹ Brinton (1981: 173) translates the above as ‘his face changed (NOT began to change) very often before he might see the chapel’; I have provided an alternative translation.

occurring in these verbs when they are auxiliated”. Since subcategorisation for the plain infinitive is a formal property of auxiliary verbs, the occurrence of *gan* and *can* with this type of infinitive will also be used to determine the extent of grammaticalization, as discussed below.

2.3.2 Syntactic Criteria

As discussed in Section 1.2, one characteristic of whether verbs are classed as auxiliary verbs is their subcategorisation for the plain infinitive, not the *to*-infinitive (Warner 1993: 136-140). In OE, both a plain infinitive and a *to*-infinitive were used, although the former was much more frequently applied than in ModE. In fact, Warner (1993: 135) argues that in OE, the predecessors of modal (auxiliary) verbs already shared a unique property in the regularity with which they were found with the plain infinitive. This is in contrast to the variation in the infinitive found with other verbs, especially those sharing the same kind of structural semantics. In relation to OE *onginnan* and OE and eME *aginnan*, Warner (1993: 137) notes, “[w]ith only two verbs is there any sign of a restriction to a plain infinitive in the complement, though the preference for the plain infinitive complement shown by *onginnan* (and *aginnan*) is notable”. Similarly, Los (2000) argues that whether the plain infinitive or the *to*-infinitive is used with OE *onginnan* and OE *beginnan* determines the extent of grammaticalization of these verbs. Warner (1993: 139) continues that the plain infinitive continued to be displaced by the *to*-infinitive, except for auxiliary verbs.⁴⁰ However, with the verb *gan*, it appears that variation in relation to the choice of infinitive became the norm in ME, as in (12) below and (13) below (*MED* s.v. *ginnen* v. 3b, *OED* s.v. *gin* v.¹ 1a, Beschorner 1920: 9-11,

⁴⁰ This excludes a number of cases: a) the plain infinitive does not always occur after non-finites, such as the modal (auxiliary) verb *may*, which can also be found with the *to*-infinitive; b) the *to*-infinitive sporadically occurs in the majority of cases when a complement infinitive is separated from the main verb, as in the case of co-ordination, where the subsequent infinitive may have *to*, but not when directly following its conjunction. There are also a small number of cases where the material that is not part of co-ordination intervenes between the main verb and the infinitive. Warner (1993: 138) argues that “[t]his all seems to be part of a general Middle English tendency to stronger infinitive marking (or perhaps a tolerance of less constrained infinitive marking) in complement infinitives separated from their governing verbs”; and c) the modal (auxiliary) verb *will* occurs with the *to*-infinitive in the 15th century (Visser 1969: §1730). However, Warner (1993: 138) argues that in ME, the modal (auxiliary) verb *will* and its preterite form *would* (*OED* s.v. *will* v.¹) could be confused with the lexical verb *will* and its preterite form *willed* ‘to wish, desire; sometimes with implication of intention’ (*OED* s.v. *will* v.²), especially that both verbs could be difficult to distinguish, as noted by the *OED* (s.v. *will* v.¹ 48).

Mustanoja 1960: 531, Visser 1969: 1373-1375, 1379-1380, 1571-1576, Brinton 1981: 156; 1988: 120, Ogura 1997: 409, Sims 2008: 140-151). Warner (1993: 139) observes the following,

[i]nterestingly even uses of *ginnen* which have been classified as ‘auxiliary’ on semantic grounds seem to lack a restriction to the plain infinitive. The complement of aspectual *ginnen* varies (see *MED ginnen* v. 3a and examples cited in Brinton 1988). *MED* isolates meaningless ‘auxiliary’ uses. But even here there is variation between the *to*-infinitive and the plain infinitive (*MED ginnen* v. 3b). Middle English *ginnen* may well show a preference for the plain infinitive, especially in weakened senses. But it occurs not infrequently with the *to*-infinitive, and it has yet to be demonstrated that any particular subset of its uses is restricted to the plain infinitive.

- (12) *He gann þennkenn off himm self*
 He did/began to think of him self
 (?c1200 *Ormulum* (Burchfield Transcript) l. 3274; *OED* s.v. *gin* v.¹ 1a)

‘He did/began to think about himself’

- (13) *His heorte gon to wakien*
 His heart began to awake
 (Layamon 1979: 7; Visser 1969: 1379, as cited in Brinton 1983: 239; translation in the original)

‘His heart began to awake’

This same kind of variation in the choice of infinitive is also reported for the verb *gan* in Scots. For example, Moessner (1997: 114, cf. *DSL* s.v. *gan* p.t.) notes,

[a]n expansion by *gin* (present) or *gan* (past) can be followed by the bare infinitive or by the *to*-infinitive. Although formally similar to the expansion by a modal auxiliary, it is syntactically and semantically different; syntactically, because it does not combine with other expansions, semantically, because *gin/gan* does not add a modal component.

Like the verb *gan*, the variant *can* displays variation in ME in terms of its infinitival complement. The *MED* (s.v. *can* v. a) and the *OED* (s.v. *can* v.² I 2) show that when it occurs with the *to*-infinitive (rarely), it has the meaning of ‘began, fell, proceed to’, as in (14) below, whereas with the plain infinitive, the meaning is ‘approached or passed into a simple auxiliary of the past tense = the modern *did*’ (*MED* s.v. *can* v. a), as in (15) below. However, in Scots, the variant *can* is almost exclusively used

with the verb in the plain infinitive, in which case it occurs with the non-ingressive sense, as in (16) below (*DSL* s.v. *can* v. 2a), with the *OED* (s.v. *can* v.² I 2) only recording two such instances. In contrast, the variant *couth*, as in (17) below, is found with the plain infinitive only (*DSL* s.v. *can* v. 2b, *OED* s.v. *can* v.² II).

- (14) *Fast þai can* [Göttingen MS Theol. 107 *gan*, Oxford, Bodleian Library, Fairfax MS 14 *con*, Trinity College, Cambridge MS R.3.8 *gon*]
 Fast they began
on him to stare
 on him to stare
 (a1400 (a1325) *Cursor Mundi* (Cotton Vespasian A.3 MS) 13557;
OED s.v. *can* v.² I 2)

‘They began to stare at him intently’

- (15) *Moyses on þe roche kan stand*
 Moses on the rock did stand
 (a1400 (a1325) *Cursor Mundi* (Cotton Vespasian A.3 MS) 6390;
MED s.v. *can* v. a)

‘Moses stood on the rock’

- (16) *And sone to Parys can he ga*
 And before long to Paris did he go
 (1375 John Barbour *The Bruce* I 330; *DSL* s.v. *can* v. 2a)

‘And he set forth for Paris soon afterwards’

- (17) *Ay he grate, quhen he cuth mene*
 Ever he wept, when he did think
Of his master swet spekinge
 Of his master sweet speaking
 (1380 *The Scottish Legendary*; *DSL* s.v. *can* v. 2a)

‘He continued weeping, when he thought of his master’s beautiful manner of speaking’

In relation to their complements, therefore, the above overview shows that in ME, *gan* and *can* favour, but are not restricted to, the plain infinitive, especially when this verb and its variant are used in a non-ingressive sense. Furthermore, in Scots, the variant *can*, including its preterite counterpart *couth*, expresses a stronger preference for the plain infinitive than the verb *gan*. These two last points raise an important

issue about the nature and processes of grammaticalization. It is commonly thought in the literature that the more grammaticalized variant shows a preference for the plain infinitive (Warner 1993: 103, Los 1998; 2000; 2005: 82-85, 88-99). However, a single factor involving the occurrence of *gan* and *can* with the plain infinitive is not enough to determine whether the above have or have not undergone grammaticalization – we can only be certain if other processes, for example semantic bleaching, have taken place, too. Overall, therefore, applying the diagnostic that the more grammaticalized the verb, the more likely to occur with the plain infinitive, together with the semantic diagnostics, as outlined in the previous section, can help isolate instances of *gan* and *can* in auxiliary use in the NME and ESc data sets. In what follows, I identify specific methods used towards identifying instances of this verb and its variant in these data sets.

2.4 Research Approaches Used Towards Identifying Items under Investigation

An obvious advantage of using a corpus is linked to the fact that a corpus can provide quantitative data. However, having this kind of data cannot explain the complex relationship between linguistic and contextual factors. Therefore, quantitative data should be complemented with qualitative data in a corpus study. There are differences in the methods used to elicit both types of data: the quantitative method is applied to examine the frequencies of occurrence of the linguistic items under investigation, with a view to identify their patterns of variation and/or change; the qualitative method, on the other hand, involves investigating linguistic features in order to provide a description of their usage, taking into account other elements, such as the type of text and the context. In what follows, I discuss both research methods; I begin by looking at quantitative method, paying special attention to methodological issues associated with the quantitative research, especially in relation to the context of grammaticalization.

2.4.1 Quantitative Method

Quantitative method allows comparisons of data between different corpora using techniques appropriate to obtaining such data. McEnery & Wilson (2001: 76)

maintain that this kind of method “enables one to discover which phenomena are likely to be genuine reflections of the behaviour of a language or variety and which are merely chance occurrences.” The most straightforward way to working with quantitative data is by classifying items according to a specific scheme and carrying out an arithmetical count of the tokens of an item within the material compared, i.e. data sets, periods, texts, samples etc. In this study, the frequencies of occurrence of the verb *gan* and the variant *can*, as well as the verb *beginnen*, i.e. dependent variables, were, therefore, obtained in relation to: a) the inflectional morphology of these verbs, i.e. inflectional endings and/or changes in the root vowel to form finite and non-finite forms, where applicable; b) the syntactic properties of these verbs, i.e. the type(s) of complementation (plain infinitive vs. other types of infinitive; absolute, transitive, intransitive uses of these verbs); c) the semantics of these verbs (non-ingressive, ingressive and/or uncategorised); d) the occurrence of these verbs in prose and/or verse, and the genre of a text in which these verbs are either more or less frequently found; and e) the occurrence of these verbs in context, more specifically, whether or not the verb in infinitive is found in rhyme position. On the basis of the frequencies of occurrence uncovered, emerging patterns can, then, be interpreted and analysed. Overall, the main advantage of this research method is that the researcher is more objective about the findings of their research. This is, therefore, achieved by: firstly, measuring the data by counting the tokens of the verb *gan* and its variant *can*, as well as *beginnen*, and their complementation; and secondly, analysing the data, for instance, to show that unlike in the case of *beginnen*, the distribution of *gan* and *can* is skewed in NME and ESc texts or that this verb and its variant are used in verse only, or that their distribution dominates in narrative texts.

However, there are a number of drawbacks associated with quantitative method. Firstly, although usually an advantage, this method denotes precise categorisation, instead of ‘fuzzy categories’ – employing precise categorisation could be particularly problematic in relation to the meaning of *gan* and *can*, especially that the distinctions between this verb and its variant in this respect can be “largely arbitrary” in a number of cases (Visser 1969: 1379, cf. Einkenkel 1891, Taylor 1917: 55, Kerkhof 1966: 158-

159, Brinton 1981: 159, Denison 1993: 322). However, this issue could potentially be resolved by a systematic application of the semantic criteria outlined in Section 2.3.1 to all instances of *gan* and *can*. Next, quantitative method may favour only common occurrences of linguistic features under investigation, disregarding the less frequent ones. For example, the results of Taylor's (1917) and Terasawa's (1974) studies show that the variant *couth* is less frequently found in OSc texts than the variant *can*, even though the *OED* (s.v. *can* v.²) argues that "[i]n Older Scots, forms of the past tense of *can* [as a modal (auxiliary) verb, namely] *couth*, *cowd* and *could* ... frequently replaced the original *can* [as a variant of *gan*] in its function as a periphrastic auxiliary of the past tense". Finally, the main disadvantage of this research method is that the context in which *gan* and *can* occur is largely ignored, especially that this verb and its variant perform a function or functions in specific texts, namely narrative ones (Wuth 1915, Taylor 1917: 574, Beschorner 1920: 15-19, Funke 1922: 8, 22, Häusermann 1930: 18-22, Koziol 1932: 132-133, Homann 1954: 389-398, Mustanoja 1960: 611-614; 1983: 59-64, Kerkhof 1966: 154-159, Smyser 1967: 68-69, Visser 1969: 1572, Terasawa 1974: 101-104, Tajima 1975: 434, 437-438, Brinton 1981: 164; 1996: 67-83, Richardson 1991a: 76-89, Fischer 1992: 265-267, Ogura 1998: 300, Sims 2008: 140, Fulk 2012: 106-107). Despite the drawbacks, quantitative method is useful in that it provides the researcher with an unbiased overview of the distribution of an item in a particular variety, text etc.

Furthermore, quantification in corpus linguistics often goes beyond the simple counting of frequencies of occurrence of instances of the verb *gan* and its variant *can*, as well as the verb *beginnen* in the NME and ESc data sets. Statistical techniques can be employed to provide a mathematical analysis of the data to establish a degree of certainty between the items under investigation and the size of the material compared. Although there are a number of statistical techniques available, I concentrate here on the ones that are relevant to this study. Firstly, in all tables containing frequency counts related to *gan* and *can*, as well as *beginnen* in the NME and ESc data sets, I provided the arithmetic count of the number of these verbs (i.e. tokens). Although it is possible to use absolute (raw) frequencies where no comparisons are made between various types of the material compared, i.e. data sets,

periods, texts, samples etc. (independent variables), a higher observed absolute frequency (parameter N) of occurrence of *gan* and *can*, as well as *beginnen* in data sets, texts etc. does not necessarily mean that these verbs are most frequent, since these types of frequencies simply denote the number of instances in the material compared (parameter T^{41}). As this study aims to provide comparisons between distributions of *gan* and *can*, as well as *beginnen*, including their complements, in the data sets, periods and/or texts, I, therefore, provided observed relative frequencies (parameter R), given that the data sets, periods, texts, samples etc. (i.e. independent variables) are of markedly different sizes, as shown, for example, in Table 4. These relative frequencies were calculated by dividing the observed absolute frequency by the word count of the material compared and multiplying the result by the common base for normalisation, as in Equation 1.⁴² However, the common base for normalisation must not be at odds with the size of the corpora under investigation, as normalising frequencies per a small number of words when the corpora used here are quite considerable in size would mean that the results would look unnaturally reduced. For this reason, I normalised the observed relative frequencies per 100,000 words.

⁴¹ However, I have made adjustments for the fact that the amount of material in the NME and ESc data sets is of different size (this refers to the data sets, periods, texts and the types of text or their mode). Therefore, I calculated the observed relative frequencies using different values for parameter T , depending on the type of material with which I dealt: a) parameter R_1 – the observed relative frequency counts have been normalised using the number of words (parameter T_1) of the samples of text within a period in the NME and ESc data sets; b) parameter R_2 – the observed relative frequency counts have been normalised using the number of words (parameter T_2) of the sample of a text from the NME and ESc data sets, as outlined in Appendix One; and c) parameter R_3 – the observed relative frequency counts have been normalised using the number of words (parameter T_3) of the sample of a verse text from the NME and ESc data sets (this is due to the fact that *gan* and *can* are recorded in verse or verse written as prose only in the NME and ESc data sets).

⁴² I only provided relative frequencies alongside the absolute (raw) frequencies of occurrence of *gan* and *can*, as well as *beginnen* in this chapter and in Chapter Six. I would argue that providing relative frequencies was not relevant for the discussion of morphological forms of the above verbs in the NME and ESc data sets, as discussed in Chapter Three, types of complementation with which the above verbs are found in the above-mentioned data sets, as outlined in Chapter Four, or different types of meaning characterising these verbs in the data sets under investigation, as discussed in Chapter Five.

Table 4

Word Count of Prose and Verse Texts in the NME and ESc Data Sets in Period I and Period II

	NME Data Set	ESc Data Set
Period I	109,560	33,279
Period II	284,346	468,000

Equation 1

*Equation for the Observed Relative Frequency of **gan** and **can**, as well as **beginnen** in the NME and ESc Data Sets in Period I and Period II*

$$R = \frac{\text{observed absolute frequency (= parameter N)} \times \text{common base for normalisation (=100,000 words)}}{\text{word count (=parameter T)}}$$

Although observed relative frequencies of **gan** and **can**, as well as **beginnen** can allow comparisons across the data sets, periods, texts, samples etc., there is always a possibility that any differences between them have arisen by chance, especially that there is variation in the data sets (dialectal, lexical, grammatical, stylistic, socio-linguistic etc.). For this reason, I employed a technique for testing how statistically significant are the correlations or differences between two types of variables, i.e. dependent variables (**gan** and **can** as well as **beginnen**) and independent variables (data sets, periods, texts, samples etc.). Having the outcome of the statistical test enabled me to assert with a degree of confidence whether or not the results of my analysis are significant. Using the statistical test also gave me an advantage of being able to identify significant differences from those that are not so significant.

Although there are a number of statistical tests, I applied the log-likelihood test to carry out the statistical calculations (Rayson 1993-2014; M. Rutkowska, personal communication, 24 June 2017). Although similar to the chi-square test, there is a notable difference between both: the chi-square test presupposes a so-called normal

distribution of the data, whereas the log-likelihood test makes no such assumptions about it (Dunning 1993: 65-66, Oakes 1998, McEnery, Xiao & Tono 2006: 55, M. Rutkowska, personal communication, 24 June 2017); therefore, that no assumptions are made about the normal distribution of language data is especially important in relation to word frequencies (McEnery & Hardie 2012: 51-52). So, given that the distribution of *gan* is “skewed” in ME (Brinton 1996: 69) and that the material compared is diverse in terms of size, it would be natural to test the hypothesis whether or not the distribution of this verb and its variant varies across the material compared, and whether or not the difference in distribution, if any, is statistically significant.

2.4.1.1 Quantitative Analysis

2.4.1.1.1 *Gan and Can*

In reference to *gan* and *can*, the literature on the subject also provides some descriptions of frequencies of occurrence of this verb and its variant in ME and/or Scots. However, there is a degree of variation in relation to the amount of detail about the frequency of occurrence of this verb and its variant in these varieties. For example, Einkenkel (1891: 89) refers to *gan* as “das beliebteste periphrasticum des Me [the most beloved periphrase of Middle English]”, suggesting, relatively speaking, its frequent use, whereas Brinton (1981: 156), quoting Funke (1922: 7) and Beschorner (1920: 9-10), comments, “[a]lthough not uncommon in Old English and the other Germanic languages, in Middle English [ginnen + infinitive] is very widespread ... if somewhat random in its appearance”, suggesting diatopic spread, rather than the high frequency of occurrence. Other sources dealing with *gan* and *can* like the *OED* (s.v. *gin* v.¹ 1a), Wuth (1915: 54-57), Funke (1922), Häusermann (1930: 18-20), Mustanoja (1960: 610-615), Visser (1969: 1379-1380, 1571-1575), Brinton (1983: 235; 1988: 117-122) and Richardson (1991a: 165-166) also make statements, but often in passing, about how frequently this verb and its variant are found in ME texts. For example, the *OED* (s.v. *gin* v.¹ 1a) claims that “[*gan*] was commonly used in a weakened sense”, Brinton (1983: 235) maintains that “[this verb] has frequent and widespread use during [ME]”, whereas Richardson (1991a: 165-166) finds that *gan* with infinitive is “common” in *The Book of the Duchess* and

The House of Fame and “very common” in *Troilus and Criseyde*, but “relatively rare” in *Canterbury Tales*, “line for line, gan is used about half as often in the Miller’s Tale as in the Troilus”. Such general statements about how frequently **gan** and **can** are used in ME are probably due to the fact that these sources are more interested in the qualitative analysis of this verb and its variant, more specifically their grammatical status, meaning and/or function.

However, sources do also provide more detailed frequencies of occurrence of **gan** and **can** in the data sets that they use, or refer to, although most of the time, such frequencies are absolute (raw), i.e. consisting of the actual count of tokens.

Consequently, it is difficult to interpret the extent of the distribution of this verb and its variant in ME and Scots texts on the basis of such statements. Furthermore, the absolute (raw) frequencies must also be treated with caution, as they should only be used where no comparisons between corpora or texts are made. For example, Taylor (1917: 574) finds 189 instances of **gan** in Barbour’s *The Bruce*, but 281 instances of **can** in Blind Harry’s *Wallace*, without providing the lengths of the samples used.⁴³ Next is Koziol (1932: 132-133), who identifies 22 instances of **gan** with infinitive in *Sir Gawain and the Green Knight* and 54 such cases in *Pearl*, but very few in *Purity and Patience*, whereas Richardson (1991a: 84) finds instances of **can** in *Sir Gawain and the Green Knight*, though “far fewer uses of ingressive constructions in Gawain than in Troilus (20 compared to 302)”. Detailed absolute (raw) frequencies of occurrence of **gan** and **can** in a wide range of dialectal texts, segregated into genre type and period can also be found in a study by Terasawa (1974: 94-96), whereas Brinton (1996: 75) looks into the grammaticalization of **gan**, more specifically, its textual and interpersonal functions in a single text by Chaucer, namely *Troilus and*

⁴³ From Taylor’s (1917: 584) study, it appears that **can** is more frequent than **couth** in ESc and MSc. He finds the following proportions of **can** and **couth** in ESc and MSc: Barbour’s *The Bruce* has 185 instances of **can** and four of **couth**; Wyntoun’s *The Original Chronicle of Scotland* 69 cases of **can** and 26 of **couth**; Blind Harry’s *Wallace* stands with 138 incidences of **can** and 145 of **couth**; Holland’s *The Book of the Howlat* features one instance of **can** and 16 of **couth** 16; and Henryson’s poems have 66 instances of **can** and 57 of **couth**. He also looks at the distribution of **can** and **couth** in other ESc and MSc works, but only to ascertain their authorship using the above findings. However, some of the results of the tests he carries out do not add up – for Wyntoun’s *The Original Chronicle of Scotland*, Taylor (1917: 578, 584) argues, “[i]n all, 105 cases of the idiom were observed ... but in 12 of these the meaning of **couth** is doubtful. If these cases are omitted from the reckoning, the score stands 69 in favour of **can** to 26 in favour of **couth**”. Taylor (1917) also uses observed absolute frequencies to compare the results.

Criseyde, but still supplementing her analysis with frequencies of occurrence of this verb there, “314 examples of *gan*, *gonne*, or *gonnen* in 8239 lines of text”.

Despite having only absolute (raw) frequencies of occurrence of *gan* and *can*, the findings of the above-mentioned studies, especially those by Terasawa (1974: 94-96), suggest that the distribution of this verb and its variant in ME and Scots texts may be variety-, author-, literary work- and/or genre/type of text-specific. In particular, Brinton (1996) justifies the investigation of *gan* in Chaucer’s *Troilus and Criseyde* because this text is characterised by the highest frequency of *gan* in it. However, like the results from the various sources outlined in the previous paragraph, the results quoted by Brinton (1996: 69) are difficult to quantify, too, ranging from “[this verb being] rare in prose and not at all frequent in unrhymed alliterative verse” (Funke 1922: 22, Koziol 1932: 131-134 and Terasawa 1974: 99-100, as cited in Brinton 1996: 69) to “Chaucer uses *gan* 669 times in his rhymed verse, he uses it only three times in the ‘The tale of Malibee’ and not at all in his other prose works” (Oizumi 1991-92 s. v.v. *gan*, *gonne*, *gonnen*, as cited in Brinton 1996: 69) to “[g]an occurs only four times in the works of Malory” (Kato 1974 s.v. *gan*, as cited in Brinton 1996: 69) to “*gan* ... reaching as high as 4.5% of the total words in *King Horn*” (Terasawa 1974: 94-96, 100, as cited in Brinton 1996: 69). Overall, therefore, some studies dealing with, or referring to, the distribution of *gan* and *can* in ME and Scots provide absolute (raw) frequencies, rather than normalised ones of occurrence of this verb and its variant, even though they make comparisons in relation to this verb and its variant across corpora or their segments.

There are a limited number of sources which use, or enable the analyst to perform, some sort of normalised frequency counts of *gan* and *can* in ME and Scots texts, thus allowing more objective comparisons in terms of the actual distribution of this verb and its variant. For example, still leaving normalised frequencies to work out by the reader, Homann (1954: 394) lists the number of instances of *gan* per sample: 161 occurrences of this verb in Chaucer’s *Canterbury Tales* (19,435 lines), 289 instances in *Troilus and Criseyde* (8,239 lines), with Gower’s *Confessio Amantis* (28,708 lines) and *Pearl* having 60 instances (1,212 lines), and *Sir Gawain and the Green*

Knicht 20 instances (505) lines. Tajima (1975: 432-433) goes a step further and normalises Homann's (1954: 394) "[a] few statistics" to a common base, namely, "per hundred lines of verses as follows: *Canterbury Tales* 0.83; *Troilus and Criseyde* 3.51; *Confessio Amantis* 0.33; *Pearl* 4.95; *Gawain* 3.96". Next is Smyser (1967: 83)⁴⁴, who observes that, "*gan* appears per hundred lines of verses as follows: [*Book of the Duchess*] .97; [*Parliament of Fowles*] 3.29; [*House of Fame*] 3.66; [*Troilus and Criseyde*] 3.93; [*Legend of Good Women*] 1.83; and the *Canterbury Tales* (all verses) .98" (Smyser 1967: 83). From these accounts, the following picture emerges: frequencies of occurrence of *gan* and *can* in the above studies are carried out per lines of text, rather than words in a text; this constitutes a reasonable strategy to adopt, especially if no immediate access is available to word counts. However, despite the availability of computerised *Ormulum*, *Brut* and *King Horn*, Sims (2008: 141-144) also reports her findings using non-normalised counts per lines of text, "[i]n the approximately 20,000 half-lines of *Ormulum*, there is one example of *onginnan* alongside two examples of *gan* (and approximately forty-six occurrence of *beginnen*)." "In the 16,095 long-lines of the Caligula version [of *Lazamon's Brut*,] I count approximately 380 examples of *gan/gunnen* (and thirty-six examples of periphrastic *beginnen*)." "[In *King Horn*] we see that in the 1,650 half-lines there are approximately seventy *gan* occurrences."

In the light of the points made above, it appeared to me that my findings on frequencies of occurrence of *gan* and *can* in the NME and ESc data sets would be of some value. These are presented in Table 5 and Table 6, and show normalised frequencies of occurrence of these linguistic items under investigation, in addition to observed absolute frequencies, in order to be able to make comparisons in relation to the distribution of this verb and its variant across the data sets and periods.

⁴⁴ There are some minor differences between Homann's (1954) and Smyser's (1967) frequency counts in relation to *Canterbury Tales* and *Troilus and Criseyde*, probably because the former takes into account *gan* and *can*, whereas the latter *can* only.

Table 5

*Frequency of Occurrence of **gan** and **can** with Infinitive in the NME Data Set in Period I and Period II⁴⁵*

	gan			can		
	<i>N</i>	<i>R₁</i>	%	<i>N</i>	<i>R₁</i>	%
Period I	17	16	68	8	7	32
Period II	75	26	86	12	4	14

*Notes: N – Observed absolute frequency
R₁ – Observed relative frequency per period per 100,000 words
% – Proportion of **gan** and **can** to each other per period
NB: Observed relative frequency is rounded to the nearest integer*

Table 6

*Frequency of Occurrence of **gan** and **can** with Infinitive in the ESc Data Set in Period I and Period II*

	gan			can		
	<i>N</i>	<i>R₁</i>	%	<i>N</i>	<i>R₁</i>	%
Period I	-	-	-	8	24	100
Period II	-	-	-	-	-	-

*Notes: N – Observed absolute frequency
R₁ – Observed relative frequency per period per 100,000 words
% – Proportion of **gan** and **can** to each other per period
NB: Observed relative frequency is rounded to the nearest integer*

It can be seen from Table 5 and Table 6 that **gan** is only recorded in the NME data set in Period I and Period II, as in (18) below and (19) below, respectively, with no instances of this verb in the ESc data set (the variant **can** is dealt with below). This appears to be in line with statements in Macafee (1992/93: §8.13), who notes that in Scots, most poets use **can**, except for Barbour and Douglas, whose works are not

⁴⁵ Due to the fact that every per-cent of occurrence of **gan** and **can**, as well as **beginnen** in the tables used in this study is rounded to the nearest integer, in some tables, these values do not always add up to 100%, but may return either 99% or 101% in some cases instead.

included in this study. (However, Taylor 1917: 574 implies that the distribution of *gan* in Barbour’s *The Bruce* may be MS-specific, “[i]n the Edinburgh [MS] of Barbour’s *Bruce* ... the form *gan* is used almost exclusively throughout ... [whereas in the Cambridge MS] *can* is used consistently instead of *gan* throughout this [MS]”). Furthermore, the findings in the tables above also show that the frequency of occurrence of *gan* increases in the NME data set in Period II. In relation to the observed relative frequency, this increase is by a factor of 1.6 (63% increase). In terms of the statistical analysis, the null hypothesis is that there is no correlation between the differences in the distribution of *gan* in Period I and Period II in the NME data set. With the log-likelihood for this verb being 4.34, the result is, therefore, statistically significant at the level of $p < 0.05$ (95th percentile, 5% level). The probability of the results – i.e. the differences in the distribution of this verb between the periods in the NME data set – happening by chance is more 5%. Therefore, the alternative hypothesis, that there is a relationship in the frequency of *gan* between periods in the NME data set, should be accepted at the level of $p < 0.05$ (95th percentile, 5% level).

(18) *Pe woman þan gan hir advise*
 The woman then did her[self] advise
And answerd to him on þis wise
 And answered to him on this manner
 (1350-1420 *The Northern Homily Cycle* P II p. 77; *HCET*)

‘The woman then considered her judgement, and answered him in this way’

(19) *Adam, thyselffe made al þis syte,*
 Adam, thyself made all this grief,
For to the tree þou wente full tyte
 For to the tree thou went full quickly
And boldelyon the frute gan byte
 And boldly on the fruit did/began to bite
My lord forbed
 My lord forbade
 (1420-1500 *The York Plays* p. 70; *HCET*)

‘Adam, you made all this grief yourself, because you very quickly went to the tree and boldly did/began to bite on the fruit that my lord forbade’

Table 5 and Table 6 also demonstrate that unlike the verb *gan*, the variant *can* is recorded in both the NME and the ESc data sets, as in (20) below and (21) below, respectively, although in the latter data set in Period I only. Interestingly, there is a decrease in the application of this variant in the NME data set in Period II by a factor of 0.6 (43% decrease). This decrease appears to be correlated with an increase in the application of *gan* in this data set – while *can* constitutes 32% of all instances of *gan* and *can* in Period I, this becomes 14% in Period II. However, these increases and decreases in the application of *gan* and *can* should be treated with caution – it will be shown below that some texts are higher in the frequency of this verb and its variant than others, resulting in a skewed distribution of *gan* and *can* in the NME and ESc data sets (cf. Brinton 1996: 69). Next, *can* also appears to be more frequent in the ESc data set than in the NME data set in Period I, with over three times as many instances of this variant in the former. However, this is in line with Macafee's (1992/93: §8.13) observations about this variant. In terms of statistical analysis, like with the verb *gan*, the null hypothesis is that there is no correlation between the differences in frequency of the variant *can* in Period I and Period II in both data sets, with the alternative hypothesis being that there is such a correlation. With the log-likelihood for *can* in the NME data set being 1.38 and in the ESc data set this value being 43.40, the difference in frequency of this variant between Period I and Period II is not statistically significant in relation to this variant between periods in the former data set at the p value being < 0.05 (95th percentile, 5% level), but it is statistically significant between periods in the latter data set at $p > 0.0001$ (99.99th percentile, 0.01% level). Therefore, we should accept the null hypothesis in the former case and reject it in the latter one. However, it is also essential to establish that statistically, there is no relationship between the differences in frequency of *can* in the NME and ESc data sets (null hypothesis) – the log-likelihood for this variant is in this case is 8.61 and so we should reject the null hypothesis at the level of $p > 0.01$ (99th percentile, 1% level).

- (20) *Pis herd dauid, and forth can stand,*
 This heard David, and forth did stand,
“Sir king,” he said, “hald me couenand,
“Sir king,” he said, “hold me covenant,
I tru truli in godds might

I trust truly in God's might
(1350-1420 *Cursor Mundi* p. 432; *HCET*)

'David heard this, and stepped forward, "Sir king," he said, "keep
promise to me, I truly have faith in God's might'

(21) *Sancte Julyane than and his wyf*
Saint Julian then and his wife
To God led thai sa thankful lyfe –
To God led they so thankful life –
That thare dedis sa wele can stere
That their deeds so well did guide
The lytil tyme that tha lifit here –
The little time that they lived here –
(1375-1420 *The Life of St. Julian Hospitaller* l. 473-476; *TEAMS*)

'Saint Julian and his wife then led a life pleasing to God, who guided
their actions well during the short time they lived here'

It is naturally desirable to check the validity of the statements made by the sources referred to at the beginning of this section about the frequencies of occurrence of *gan* and *can* in their textual evidence against the data sets used in this study (even though I do not apply statistical analysis here, I still quantify with relation to the text type). The distribution of *gan* and *can* in the NME and ESc data sets, as presented in Table 7 and Table 8, suggests that the occurrence of this verb and its variant may be specific to verse texts, with only two instances of *gan* being recorded in what the editors of the *HCET* class as a prose text (cf. Smyser 1967: 69, Terasawa 1974: 94-96, Brinton 1981: 164; 1988: 120; 1996: 69). However, as will become apparent in Chapter Six, the two instances of *gan* in a prose text are, in fact, written in verse and presented as prose. Taking all this into consideration, my findings appear to be similar to those found in Visser (1969: 1572) and Fulk (2012: 106), who argue that the use of this verb and its variant is restricted to verse exclusively. Furthermore, given that *gan* and *can* is either exclusively or predominantly found in verse, Taylor (1917: 574), Beschorner (1920: 15-19), Funke (1922: 22), Koziol (1932: 132-133), Mustanoja (1960: 610-615; 1983), Kerkhof (1966: 154-159), Smyser (1967: 68-69), Visser (1969: 1572), Terasawa (1974: 102), Tajima (1975: 434, 437-438), Richardson (1991a: 80-81), Fischer (1992: 266), Ogura (1998: 300), Sims (2008: 140) and Fulk (2012: 106-107) argue that this verb and its variant, where references

are made to *can* in these sources, have a function or functions there (the distribution and function of *gan* and *can* in NME and ESc verse will be addressed in Chapter Six).

Table 7

Occurrence of gan and can with Infinitive in Texts in the NME Data Set in Period I and Period II

	Text	<i>gan</i>		<i>can</i>	
		<i>N</i>	<i>R</i> ₂	<i>N</i>	<i>R</i> ₂
Period I	<i>Cursor Mundi (HCET)</i>	2	19	8	76
	<i>The Northern Homily Cycle (HCET)</i>	12	158	-	-
	<i>The Pricke of Conscience (HCET)</i>	1	17	-	-
	<i>Rolle's Prose Treatises (ICoMEP)</i>	2	2	-	-
Period II	<i>Sir Eglamour of Artois (TEAMS)</i>	18	208	1	12
	<i>Octavian (TEAMS)</i>	56	487	11	96
	<i>The York Plays (HCET)</i>	1	18	-	-

Notes: *N* – Observed absolute frequency
*R*₂ – Observed relative frequency per text per 100,000 words
 NB: Observed relative frequency is rounded to the nearest integer

Table 8

Occurrence of can with Infinitive in Texts in the ESc Data Set in Period I

	Text	<i>can</i>	
		<i>N</i>	<i>R</i> ₂
Period I	<i>St. Andrew and the Three Questions (TEAMS)</i>	2	109
	<i>The Life of St. Julian Hospitaller (TEAMS)</i>	6	320

Notes: *N* – Observed absolute frequency
*R*₂ – Observed relative frequency per text per 100,000 words

As already mentioned above, the use of **gan** and **can** in verse has been linked to a specific type of text, with the findings from both Terasawa (1974: 94-96) and Brinton (1996: 69) clearly demonstrating this. My findings show that **gan** occurs in the highest numbers (observed relative frequency) in *Octavian*, in comparison to the other texts used in the NME data set. Furthermore, samples from *Sir Eglamour of Artois*, *The Northern Homily Cycle* and *Cursor Mundi* have a considerably ‘high’ distribution of this verb, too, in comparison to Rolle’s *Prose Treatises*, *The Pricke of Conscience* and *The York Plays*. In relation to **can** in the same data sets, *Octavian* seems to have the highest relative frequency of this variant, although next is *Cursor Mundi*. In the ESc data set, on the other hand, the relative frequencies of **can** show that both *The Life of St. Julian Hospitaller* and *St. Andrew and the Three Questions* are also ‘high’ in the occurrence of this variant, although the former leads the way with nearly three times as many instances than the latter, when considered per 100,000 words. (However, the difference in observed absolute frequencies and word counts should be observed for *Octavian* in the NME data set and *The Life of St. Julian Hospitaller* in the ESc data set). Overall, the above suggests that the distribution of **gan** and **can** in texts in the NME and ESc data sets is distorted, with some texts having indeed a ‘higher’ distribution of either or both of this verb and its variant than others.

2.4.1.1.2 *Beginnen*

The verb **beginnen** and its frequency of occurrence in ME are only occasionally mentioned in the literature, usually alongside discussions on **gan** and **can**, and their etymologically related variants, or when other verbs of ingressive aspect are discussed. For example, the *OED* (s.v. *begin* v.¹) argues that “[*b*]eginnan was very rare in Old English, where the ordinary word was *onginnan*”. In a similar way, Visser (1969: 1373) suggests that this verb was also infrequently found in OE, “[s]oon after [the times of Ælfric] it began to occur with increasing frequency, with the result that by the beginning of the fifteenth century it had entirely ousted the older forms [namely OE *onginnan* and OE and eME *aginnan*].” While neither of the above sources provide arithmetic counts, Los (2000: 252) demonstrates that OE

beginnan was already frequently used in Ælfric’s works, on a par with OE *onginnan*, “[o]nginnan was complemented by the bare infinitive 44 times, and by the *to*-infinitive 22 times. *Beginnan* was complemented 13 times by a bare infinitive, and 46 times by a *to*-infinitive”. In ME, *beginnen* also appears to be frequently used, although perhaps even more so than in OE. For example, Brinton (1981: 119-120) argues that *beginnen* is “very common”, whereas Sims (2008: 133) maintains that it is “firmly established” in this variety. In Scots, this verb is found, too (*DSL* s.v. *begin* v.), but no sources deal with frequencies of this verb or, especially, its variant *begouth* in this variety. It is, therefore, difficult to know the extent to which the frequency of occurrence of *begouth* is similar to that of *beginnen* in Scots.

Given that *beginnen* is frequent in ME, as outlined by the sources quoted in the preceding paragraph, it is, therefore, unsurprising to find instances of this verb in the NME data set. Instances of this verb are also recorded in the ESc data set, with the distribution of this verb as an independent verb being similar to that of this verb in this function in the NME data set, but with a different frequency when this verb occurs with infinitive. Table 9 and Table 10, therefore, show the frequency of occurrence (absolute and relative) and the proportion (per-cent) of instances of this verb functioning as an independent verb, as in (22) below and (23) below, and with infinitival complements, as in (24) below and (25) below, in both data sets. However, the findings presented in these tables also show that overall, the verb *beginnen* (as an independent verb and with infinitive) is more frequently used in both data sets than the verb *gan* and its variant *can* are. This might be linked to the fact that this verb exhibits “historical stability” as a marker of ingressive aspect in ME already (Sims 2008: 162). In terms of statistical analysis, the null hypothesis is that there is no correlation between the differences in the distribution of *beginnen* over regions (NME and ESc), with the alternative hypothesis being that there is such a correlation. The log-likelihood value for *beginnen* as an independent verb in the NME and ESc data sets is 1.80, this suggesting that the null hypothesis should be accepted at the level of $p < 0.05$ (95th percentile, 5% level). In contrast, with the log-likelihood value for *beginnen* with infinitive being 192.56 in these data sets, the null hypothesis

should be rejected at the level of $p < 0.0001$ (99.99th percentile, 0.01% level), with the alternative hypothesis accepted.

Table 9

*Frequency of Occurrence of **beginnen** as an Independent Verb and with Infinitive in the NME Data Set in Period I and Period II*

	beginnen					
	Other			Inf		
	<i>N</i>	<i>R_I</i>	%	<i>N</i>	<i>R_I</i>	%
Period I	24	22	47	27	25	53
Period II	38	13	22	135	47	78

Notes: *N* – Observed absolute frequency
R_I – Observed relative frequency per period per 100,000 words
 % – Proportion of **beginnen** as an independent verb and with infinitive per period
 NB: Observed relative frequency is rounded to the nearest integer

Table 10

*Frequency of Occurrence of **beginnen** as an Independent Verb and with Infinitive in the ESc Data Set in Period I and Period II*

	beginnen					
	Other			Inf		
	<i>N</i>	<i>R_I</i>	%	<i>N</i>	<i>R_I</i>	%
Period I	7	21	88	1	3	13
Period II	55	12	83	11	2	17

Notes: *N* – Observed absolute frequency
R_I – Observed relative frequency per period per 100,000 words
 % – Proportion of **beginnen** as an independent verb and with infinitive per period
 NB: Observed relative frequency is rounded to the nearest integer

- (22) *Pen many begynnes þe thyng þat þai may*
 Then many begin[s] the thing that they may
neuer-mare bryng till ende
 never-more bring to end
 (1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 6;
ICoMEP)

‘Then many begin the undertaking that they may no longer bring to an end’

- (23) *EXECUCION OF ALL THINGES DON BE-TWIX*
 execution of all things done between
yAIR BOUNDES APAN THE WEST MARCHE
 their bounds upon the west boundary
BEGYNANDE yE XIIIJ DAYx OF OCTOBeR
 beginning the xiiij day of October
TO CANDELMASSs DAY LAST PASSYT
 to Candlemas day last passed
 (1385 *Rotuli Scaccarii Regum Scotorum: The Exchequer Rolls of*
Scotland 1384/85 Mar 15; LAOS)

‘execution of all things done with reference to their boundary along the western side of the border, beginning on 13 October until last Candlemas [2 Feb]’

- (24) *Pe heued biginnes for to scak,*
 The head begins for to shake,
His hend vnquemli for to quak,
 His hand unsettlingly for to tremble,
It crepes crouland in his bak
 It creeps crawling on his back
 (1350-1420 *Cursor Mundi* p. 212; *HCET*)

‘The head begins to shake, and his hand [begins] to unsettlingly tremble; it creeps crawling on his back’

- (25) *And than the wellis begynnys till habound jn*
 And then the wells begins to overflow in
wateris and stryndis to spring
 waters and streams to spring
 (1420-1500 *The Prose Works of Sir Gilbert Hay* p. 98; *HCOS*)

‘And then the wells begin overflowing and streams [begin] to flow [out of the ground]’

Looking at the frequencies of occurrence of *beginnen* on their own, it appears that this verb as an independent verb is more frequently used in the NME and ESc data

sets in Period I than in Period II, this change being by a factor of 1.7 (41% decrease) and 1.8 (43% decrease) in their respective data sets. An opposite tendency can be observed in relation to *beginnen* with infinitive, but only in the NME data set, with an increase of this construction type by a factor of 1.9 (88% increase) in Period II. With reference to the ESc data set, there is a slight decrease in the application of this construction type in Period II, by a factor of 0.7 (33% decrease). Therefore, when the distribution of this construction type is compared in both data sets, the tables show that the verb *beginnen* with infinitive displays the biggest increase in use in the NME data set in Period II. However, despite claims in the literature that *beginnen* is frequently found in ME (Brinton 1981: 119-120, Sims 2008: 133), the problem of relatively lower frequencies of occurrence in general applies to this verb in the NME and ESc data sets, too. When statistical significance tests are applied to test the null hypothesis that there is no correlation between the differences in the frequency of *beginnen* between periods in the NME data set, it appears that this is the case, i.e. we should accept the null hypothesis, for this verb as a lexical verb at the p value being < 0.05 (95th percentile, 5% level, log-likelihood = 3.43), but reject the null hypothesis for this verb with infinitive in this data set at the p value being < 0.001 (99.9th percentile, 0.01% level, log-likelihood = 11.12). Considering the same hypothesis, but for the ESc data set, it seems that the difference in the frequency of *beginnen* as an independent verb and with infinitive between Period I and Period II is not statistically significant at $p < 0.05$ (95th percentile, 5% level, log-likelihood = 1.81 and 0.05, respectively), in which cases, we should accept the null hypothesis.

The results for *beginnen* also show that in contrast to *gan* and *can*, the distribution of this verb does not appear to depend on the type of text used in the data sets in this study, as outlined in Table 11 and Table 12, with this verb being recorded in both prose and verse alike (even though I do not apply statistical analysis here, I still quantify with relation to the text type).

Table 11

Occurrence of *beginnen* as an Independent Verb and with Infinitive in Texts in the NME Data Set in Period I and Period II

		<i>beginnen</i>				
		Text	Other		Inf	
			<i>N</i>	<i>R</i> ₂	<i>N</i>	<i>R</i> ₂
Period I	<i>Cursor Mundi (HCET)</i>	1	10	10	95	
	<i>The Northern Homily Cycle (HCET)</i>	4	53	2	26	
	<i>Rolle's Yorkshire Writers (ICoMEP)</i>	19	24	15	19	
	<i>Alphabet of Tales Parts I and II (ICoMEP)</i>	14	8	94	52	
	<i>Dan Jon Gaytryge's Sermon (HCET)</i>	3	56	-	-	
	<i>Sir Eglamour of Artois (TEAMS)</i>	2	23	5	58	
	<i>The Liber de Diversis Medicinis (HCET)</i>	-	-	1	17	
	<i>Methodius (ICoMEP)</i>	2	58	6	173	
Period II	<i>The Mirror of St. Edmund (ICoMEP)</i>	2	14	8	56	
	<i>Octavian (TEAMS)</i>	-	-	9	78	
	<i>Rolle's Prose Treatises (HCET)</i>	2	32	1	16	
	<i>Rolle's The Psalter (HCET)</i>	1	8	1	8	
	<i>Rolle's Yorkshire Writers (ICoMEP)</i>	5	25	8	41	
	<i>The Wakefield Pageants (HCET)</i>	4	97	2	49	
	<i>The York Plays (HCET)</i>	3	54	-	-	

Notes: *N* – Observed absolute frequency
*R*₂ – Observed relative frequency per text per 100,000 words
 NB: Observed relative frequency is rounded to the nearest integer

Table 12

*Occurrence of **beginnen** as an Independent Verb and with Infinitive in Texts in the ESc Data Set in Period I and Period II*

		<i>beginnen</i>				
		Text	Other		Inf	
			<i>N</i>	<i>R</i> ₂	<i>N</i>	<i>R</i> ₂
Period I	<i>LAOS Texts</i> ⁴⁶	7	188	-	-	
	<i>The Life of St. Julian Hospitaller (TEAMS)</i>	-	-	1	53	
	<i>The Acts of Parliaments of Scotland (HCOS)</i>	7	26	1	4	
	<i>The Craft of Deyng (Ratis Raving) (HCOS)</i>	1	31	-	-	
	<i>De Irelandia's The Meroure of Wyssdome (HCOS)</i>	4	48	1	12	
Period II	<i>The Porteous of Noblenes (HCOS)</i>	-	-	1	27	
	<i>The Prose Works of Sir Gilbert Hay (HCOS)</i>	9	31	5	17	
	<i>The Vertewis of the Mess (Ratis Raving) (HCOS)</i>	1	122	-	-	
	<i>LAOS Texts</i>	33	229	3	363	

Notes: *N* – Observed absolute frequency

*R*₂ – Observed relative frequency per text per 100,000 words

NB: Observed relative frequency is rounded to the nearest integer

The distribution of *beginnen* in all types of texts is perhaps to be expected, as this verb continues to be used as a marker of ingressiveness throughout the history of English (Brinton 1981: 119-120, Sims 2008: 133) and in Scots (*DSL* s.v. *begin* v.), without being confined to any particular type of variety, period and/or text. In this respect, the presence of only one instance of *beginnen* with infinitive in the ESc data set in Period I overall, as in (26) below, is probably a reflection of the paucity of textual evidence from this period, rather than the availability of this verb in general at the start of the ESc period (cf. Section 1.3.2).

⁴⁶ No list of texts and their respective word counts from the *LAOS* is available in the public domain (Williamson, personal communication, 28 April 2014).

(26) *Allace! I thocht nocht fore to thryfe,*
Alas! I thought not for to thrive,
With Fortone quhen I began to strife.
With Fortune when I began to strive.
 (1375-1420 *The Life of St. Julian Hospitaller* l. 365-368; TEAMS)

‘Alas! I did not think to succeed, when I began to strive with Fortune.’

Indeed, the *DSL* (s.v. *begin* v. 1b) has the first recorded instances of ***beginnen*** with infinitive from *The Scottish Legendary* in 1380⁴⁷, but the next recorded instances being from Wyntoun’s *The Original Chronicle of Scotland* in 1420, the difference being some 40 years. Furthermore, the occurrence of this verb in different types of texts in both the NME and ESc data sets does not appear to be arrested by the rise of other such ingressive verbs at the time (Brinton 1981: 152-155).⁴⁸ Finally, although generally equally distributed in both data sets, relative frequencies of this verb appear to be higher in *Methodius*, *The Wakefield Pageants in the Towneley Cycle* and *Cursor Mundi* in the NME data set and in various local documents in the ESc data set, where it mostly occurs in the participial form referring to the time or location, as in (23) above. (The exceptions are *Extracts from the Council Register of the Burgh of Aberdeen* and *Charters and Documents Relating to the Burgh of Peebles*).

⁴⁷ Although the *DSL* provides 1380 as the date of *The Scottish Legendary*, it is written in “[t]he dialect [of] Lowland Scottish of about 1400” (Scahill & Rogerson 2005: 91).

⁴⁸ In eME and ME, according to Brinton (1981: 152-155), the following verbs take on the role as markers of ingressive aspect: ***comsen, bresten (on, out on, into, in), commencen, fallen, gon, breken on, grouen, setten, fallen (in, (up)on, a)***. In ESc, on the other hand, the *DSL* (s.vv. *brek* v. 9(c)b; *enter* v.¹ 3c; *move* v. 5, 13; *proced(e* v.; *set* v. 11, 73b; *ris(e* v. 18, 19b, 20) has ***brek, enter, procede*** and ***rise*** as having some ingressive meaning between 1375 and 1420, whilst ***move*** and ***set*** as displaying some ingressive meaning between 1420 and 1500; out of these, only ***move*** occurs with infinitive between 1375 and 1420, as in (1) below.

(1) *All off fors that hym behowyd*
All of force that him behove
As he his purpos that tyme movyde
As he his purpose that time moved
To drawe the kynryk in fredwme
To draw the country in freedom
 (c1420 Andrew of Wyntoun *The Original Chronicle of Scotland* VIII 5108; *DSL* s.v. *move* v. 13a)

‘All of force that mobilised him as he began pursuing his purpose to lead the country to freedom’

2.4.1.2 Quantitative Analysis and Grammaticalization

The above-mentioned findings show that although *beginnen* is more frequently found in the NME and ESc data sets than *gan* and *can* are, all of these verbs are relatively infrequent in the above-mentioned data sets, especially that observed relative frequencies are given here per 100,000 words. However, the frequency of occurrence for *gan* and *can*, as outlined in the previous sections, shows that some texts are higher in the occurrence of this verb and its variant than others in the NME and ESc data sets. This is, especially, the situation in relation to *Octavian* on the one end of the spectrum (observed relative frequency in this text is 487 for *gan* and 96 for *can* per 100,000 words), as *Prose Treatises* by Rolle, *The Pricke of Conscience* and *The York Plays* have all a much lower frequency of occurrence of this verb and its variant. This, in turn, suggests that in those texts where they are more frequent, i.e. almost exclusively verse texts, more specifically narrative ones, *gan* and *can* perform a function or functions (Wuth 1915, Taylor 1917: 574, Beschorner 1920: 15-19, Funke 1922: 8, 22, Häusermann 1930: 18-22, Koziol 1932: 132-133, Homann 1954: 389-398, Mustanoja 1960: 611-614; 1983: 59-64, Kerkhof 1966: 154-159, Smyser 1967: 68-69, Visser 1969: 1572, Terasawa 1974: 101-104, Tajima 1975: 434, 437-438, Brinton 1981: 164; 1996: 67-83, Richardson 1991a: 76-89, Fischer 1992: 265-267, Ogura 1998: 300, Sims 2008: 140, Fulk 2012: 106-107). The verb *beginnen* also appears to be more frequently found in some texts, for example *Methodius*, *The Wakefield Pageants in the Towneley Cycle* and *Cursor Mundi* in the NME data set and in local texts, as part of the ESc data set. However, in contrast to *gan* and *can*, the distribution of this verb is relatively equal in both prose and verse, rather than in just one or the other type of literary form.

A number of sources either argue or hint that there is a relationship between frequency of occurrence and grammaticalization in that the higher the frequency the more grammaticalized an item is (cf. also Heine, Claudi & Hiinmeyer 1991b, Bybee, Perkins & Pagliuca 1994, Haiman 1994: 9, Krug 1998: 309, Bybee 2003: 612, Hoffmann 2004: 171-210, Mair 2004: 125 and L. Wright 2004: 211-226 amongst others). For example, Hopper & Traugott ([1993] 2003: 106) maintain, “the more frequently a form occurs in texts, the more grammatical it is assumed to be.

Frequency demonstrates a kind of generalization in use patterns.” In other words, textual frequency is often considered primary evidence of degree of grammaticalization. However, I would argue that what can be considered as frequent should be assessed on an individual basis. Therefore, the lower frequency of occurrence of *gan* and *can* in this study, except for a selection of texts, can be accounted for in the following ways: a) the NME and ESc data sets do not contain enough of the verse type of texts in which this verb and its variant are likely to occur; and b) *gan* and *can* are on their way to being lost, by being replaced by periphrastic *do*, as indicated or suggested in the literature on *gan* and *can* in ME (Taylor 1917: 591, Funke 1922: 26, Mustanoja 1960: 614, Visser 1969: 1572-1573, cf. Sims 2014: 67) (the potential role of periphrastic *do* in the disappearance of *gan* and *can* is discussed in Section 2.3.1). Overall, changes in the frequency of occurrence do not reflect grammatical change directly, but may be indicative of how grammatical innovation spreads in actual usage. Therefore, when considering the relationship between the frequency of occurrence of an item and its degree of grammaticalization, the item’s usage in a specific context should also be taken into account (this will be addressed in Chapter Six).

2.4.2 Qualitative Method

Corpus linguistics is essentially quantitative. Yet, quantitative method cannot be used to account entirely for interplay of linguistic and contextual factors. Instead, qualitative method is employed to take advantage of computerised corpora to go beyond the bare statistics of occurrence of items under investigation. According to McEnery & Wilson (2001: 76), qualitative analysis allows providing a full and detailed view on the data, where both frequent and rare linguistic features have, or, at least, should have, the same amount of attention. Therefore, the purpose of the qualitative analysis is to provide a detailed account, i.e. similarities, differences and/or variation, of the findings from the data. One of the advantages of using qualitative method in this type of study would be a more detailed interpretation of the extent of grammaticalization of the verb *gan* and its variant *can*, as well as the verb *beginnen* in the NME and ESc data sets. In particular, this type of method would be used to interpret changes of the above verbs in relation to the mechanisms and

processes associated with grammaticalization: a) morphological fixation of forms of *gan* and *can* in the third person singular in the preterite (only finite forms available; root vowel changes), whereas the lack of it in relation to forms of *beginnen* (both finite and non-finite forms available; inflection and root vowel changes); b) analogy, whereby forms like *couth* and *begouth* as variant forms of *can* and *beginnen*, respectively, emerge during the process of grammaticalization; c) reanalysis, as a result of which, the verb *gan* and its variant *can* occur with the plain infinitive in view of claims in the literature that the more grammaticalized variant takes the plain infinitive, with the verb *beginnen* being found with the *to*-infinitive, as well as other types of infinitive; and d) semantic bleaching, opacity and persistence, whereby the verb *gan* and its variant *can* are found with non-ingressive or ingressive meaning, or in a number of cases, with distinctions between both meanings being difficult to make.

Other advantages of qualitative method include: a) looking further than the quantitative data. For example, in relation to the semantics of *gan* and *can*, qualitative method can be used to analyse the use of this verb and its variant in the NME and ESc data sets. On the basis of the surrounding context, i.e. the circumstances that form the setting of a particular event or situation denoted by the verb in infinitive, *gan* and *can* can be said to have either ingressive or non-ingressive meaning. Furthermore, the role of the context can play an important role when investigating the function of *gan* and *can*, namely to enable the verb in infinitive to provide rhyme and to perform a textual and interpersonal function in narrative texts written in verse. Finally, qualitative method can provide a framework for discussion of the similarities and differences amongst particular texts and genres; b) performing an inductive process, which would involve developing theories about the function *gan* and *can* in the NME and ESc data sets on the basis of the usage of this verb and its variant in the data gathered. For example, given that *gan* and *can* occur in a number of clearly identifiable contexts in verse texts in these data sets, the role of this verb and its variant can be said to be that of a narrative device. Similar observations can also be found in Brinton (1996: 77-78), who argues that in ME, the verb *gan* functions as a pragmatic marker “used in foregrounded clauses to mark

salient shifts or turns in the course of events ... *gan* serves to mark structurally significant transitions and the beginning of new events in the advancement of the plot"; and c) looking for new and emerging categories, i.e. the verb *gan* and its variant functioning as auxiliary verbs in ME and Scots, comparing and contrasting, but also identifying and interpreting deviations from patterns, such as the occurrence of *couth* and *begouth* as variant forms of *can* and *beginnen*, respectively. Overall, therefore, an important part of this corpus investigation will be to go beyond the quantitative results. This will enable functional interpretations why such patters exist and interpretations of whether the verb *gan* and its variant *can*, as well as the verb *beginnen* develop differently in terms of grammaticalization in NME and ESc.

2.5 Summary

In this chapter, I described materials and showed that in principle, using computerised corpora could be useful to investigate *gan* and *can*, as well as *beginnen* in the NME and ESc data sets, despite the fact that I essentially had to go through the data manually. I also demonstrated that despite the advantages of using computerised corpora to carry out an investigation, such sources of data might have issues with representativeness. For example, some samples may be wrongly and/or incorrectly categorised in terms of the date and/or dialect of a text. In this chapter, I also discussed what kinds of criteria should be used in the classification of items under investigation. However, I showed that, especially, in the case of *gan* and *can*, there might be problems with determining the meaning of this verb and its variant, although a combination of both semantic and syntactic criteria would be helpful to separate auxiliary uses of this verb and its variant from non-auxiliary, i.e. ingressive, ones. Finally, I provided a description of the approaches to research in this study. While employing both qualitative and quantitative methods offers a better understanding of the behaviour of linguistic items like *gan* and *can*, as well as *beginnen*, I emphasised caution in relation to the frequency of occurrence of these verbs overall. In particular, the investigation of the NME and ESc data sets showed that the frequency of *gan* and *can* appears to be skewed in the above data sets, with the distribution in both data sets being low overall, especially that relative frequencies are given per 100,000 words.

Chapter Three – Changes in Form

3.1 Introduction

This chapter considers the inflectional morphology of *gan* and *can*, as well as *beginnen* in the NME and ESc data sets. Inflectional morphology has two main functions: a) it marks grammatical categories on words; and b) it establishes a connection between elements within a sentence or discourse. Morphology as such, therefore, functions as a link, which consists of a set of complex dependencies and realisations, between both phonology and syntax, but also semantics.⁴⁹ In OE, the inflectional morphology of verbs depended on their mood, number, person and tense. During the course of ME, it is possible to observe a change that resulted in a loss of inflectional endings characterising verbs in OE. The two major types of restructuring the verb inflection concern the plural and – at a later time – the third person singular (Lass 2006: 78). Crucially for the subject matter of this study, the attrition of inflectional morphology of verbs started in the north at the end of the OE period, spreading south, finding its conclusion in the 14th and the 15th centuries (Lass 1992: 136, Fulk 2012: 71-74, Gramley 2012: 56-58). Furthermore, the attrition of inflectional morphology has also had an impact on the grammatical status of some verbs. Lightfoot (2006: 94), for example, suggests that the loss of inflection facilitated the grammaticalization of modal (auxiliary) verbs. As a result, a new class of verbs emerged – that of modal (auxiliary) verbs – characterised by the following morphological properties identified by Warner (1993: 103) to be typical of auxiliary verbs: a) the restriction of some of such verbs to finite forms only; b) the use of past-tense forms, but without referring to past time; and c) the use of preterite-present morphology.

The loss of inflectional morphology can be observed in relation to the verb *gan* and its variant *can*, as well as the verb *beginnen* in ME and in Scots. For example, both

⁴⁹ When I investigate changes in form (this chapter), complementation (Chapter Four) and meaning (Chapter Five) separately, I do not intend to imply that, especially, ‘form’ and ‘meaning’ are not interconnected, i.e. seeing ‘form’ as an autonomous entity or morphosyntactic structure and ‘meaning’ as the sense or signification of a word. On the contrary, I argue that both are linked, and so each needs to be taken into account when explaining instances of global language change affecting all verbs and grammaticalization.

Denison (1993: 322) and Brinton (1996: 81) refer to **gan** in auxiliary use as being “virtually fixed” and “[morphologically] fixed in the third-person preterite form, most often singular”, respectively, with inflected forms of this verb and its variant **can** being recorded most frequently in southern varieties of ME (*DSL* s.vv. *can* v. 2; *gan* p.t., *MED* s.vv. *can* v. I 1a; *ginnen* v., *OED* s.vv. *can* v.²; *gin* v.¹, Denison 1993: 322). The verb **beginnen**, on the other hand, is marked for mood, number, person and tense, although to a varying degree and differently in ME and in Scots (*DSL* s.v. *begin* v., *MED* s.v. *biginnen* v., *OED* s.v. *begin* v.¹, Mustanoja 1960: 610-615, Visser 1969: 1373-1375, Brinton 1981: 181-186; 1983: 241-242; 1988: 118-119, 158-159; 1996: 67, Macafee 1992/93: §8.11, Ogura 1997, Fulk 2012: 72-81). Furthermore, additional developments in relation to **can** and **beginnen** can be observed in Scots, where variant forms **couth** and **begouth** have emerged (however, Visser 1969: 1579-1580 also has instances of **couth** in NME and SME works, as discussed in Section 1.3.1) (*DSL* s.vv. *begin* v.; *can* v. 2, *OED* s.vv. *begin* v.¹; *can* v.¹, Taylor 1917, Macafee 1992/93: §8.1.1, 8.13, J. Smith 1996: 137, Markus 1997 amongst others). Given that **gan** and **can**, including the variant **couth**, occur in auxiliary use in the third person singular form in the preterite, it would appear that this verb and its variant (as well as **couth**) fulfil, at least, some of the morphological criteria for auxiliarihood outlined in the preceding paragraph.

In this chapter, I start off by a comparison with what is happening to verbs generally during the transition from OE to ME (Section 3.2). In Section 3.3, I then provide an outline of morphological forms of the verb **gan** and its variant **can**, as well as the verb **beginnen** in ME and Scots, including my own findings from the NME and ESc data sets. Firstly, I deal with the inflectional morphology of **gan** and **can**, and, then, move onto a presentation of grammatical forms of the verb **beginnen**. My findings about the morphological properties of these verbs in the NME and ESc data sets are in line with statements in the literature on the subject: a) **gan** and **can** use no inflectional suffixes to mark the distinction between the singular and plural, and are characterised by the root vowel change to form the preterite as its only form; and b) **beginnen** is, on the other hand, characterised by changes in morphology typical of other verbs at the time to form both finite and non-finite forms. This leads me to

Section 3.4, where I analyse my findings, arguing that: a) *gan* and *can* in auxiliary use are subject to decategorialization in the NME and ESc data sets, i.e. a process in which this verb and its variant lose the morphological properties that would otherwise classify both of them as full members of the lexical category of verbs; b) *begouth* arises out of the process of analogy in the ESc data set, on parallel with *couth*, which, in addition to *can*, is “[c]hiefly” found in Scots (*OED* s.v. *can* v.²), and NME and SME, too (cf. Visser 1969: 1579-1580); and c) *gan* and *can* undergo two separate processes of phonological reduction, resulting in this verb and its variant being single morphemes when in auxiliary use in the NME and ESc data sets.

3.2 Changes in the Morphology of Verbs in the Transition from Old English to Middle English

During the transition from OE to ME, the category of the verb underwent a number of radical and transforming changes in both phonology and morphology; these included: a) the attrition of inflectional morphology, which was used to convey the following verbal categories: mood (indicative, subjunctive, imperative), tense (present and past), number (singular and plural) and person (first, second, third); b) the reduction of the number of vowel grades per verb, which in the case of strong verbs were used to mark tense and number; c) the mixing of forms from other classes in the conjugation of the given verb, and transfer of verbs, either wholly or partly, into the weak conjugation (a transfer of strong verbs into the weak conjugation was common throughout the ME period, but increased in the late 14th century and the early 15th century) (Lass 1992: 131-134). By 1300, the OE system was restructured in almost all dialects of ME, but with marked differences in the north. An outline of inflections found on the verb in the north at the start of the ME period is presented in Table 13 and Table 14 for reference.⁵⁰

⁵⁰ Given the present and preterite inflections, verbs in ME were divided into: a) strong verbs; b) weak verbs; and c) irregular verbs. However, as strong verbs differed from weak ones in the preterite only, while both and strong verbs were different from irregular ones, no distinction is made here into strong and weak verbs with respect to the present indicative. Although there was a small group of verbs belonging to the former Class II of OE verbs that had a separate set of inflections in the present tense, these verbs were used in the south (Fisiak 1996: 89) and so, they are not relevant to this study.

Table 13

Inflections of the Present Indicative, Preterite, Subjunctive (Present) and Subjunctive (Preterite) of Verbs in the North in the 12th Century (Fisiak 1996: 86)

	Present indicative		Preterite		Subjunctive (present)		Subjunctive (preterite)	
	Sg	Pl	Sg	Pl	Sg	Pl	Sg	Pl
First	<∅>		<∅>					
	<e>		<d>					
	<es>		<ed>	<∅>				<∅>
	<is>		<t>	<d>				<d>
Second			<þ>	<d-e>			<∅>	<d-e>
	<es>	<∅>	<∅>	<d-en>			<d>	<d-en>
	<is>	<es>	<d-est>	<ed>	<∅>	<∅>	<d-e>	<ed>
	<s>	<is>	<ed-est>	<ed-e>	<∅>	<en>	<ed>	<ed-e>
	<st>	<s>	<t-est>	<ed-en>	<e>		<ed-e>	<ed-en>
	<t>		<þ-est>	<t>			<t>	<t>
Third			<∅>	<t-e>			<t-e>	<t-e>
	<∅>		<d>	<t-en>				<t-en>
	<es>		<ed>	<þ-en>				<þ-en>
			<t>					
		<þ>						

Table 14

Inflections of the Imperative, Present Participle, Past Participle and Infinitive of Verbs in the North in the 12th Century (Fisiak 1996: 86)

	Imperative		Present participle	Past participle	Infinitive
	Sg	Pl			
First				<d>	
Second	<∅>	<∅>	and(e)	<en>	<∅>
		<es>		<n>	
Third				<t>	
				<þ>	

The processes responsible for the restructuring of inflectional morphology on the category of the verb throughout the ME period were the same in all dialects (analogy, extension, levelling). The changes resulted in further modification to the category of the verb during this time. Given that the input in the north was different from that of the south, other developments specific of the north can be identified. What emerges

as the outcome of further changes during the ME period is the progressive disappearance of inflectional endings (as a result of analogy, extension and levelling), firstly recorded in northern dialects and, subsequently, in southern ones (Fisiak 1996: 90, Fulk 2012: 72). Inflections of the present indicative, preterite (the present indicative and the preterite of strong verbs will be dealt with below), subjunctive, imperative, present participle, past participle and infinitive of strong verbs in the north, illustrating further loss of inflectional endings, are listed in Table 15 and Table 16 for reference (all the tables contain different allomorphs existing side by side at the same time).

Table 15

Inflections of the Present Indicative, Preterite, Subjunctive (Present) and Subjunctive (Preterite) of Strong Verbs in the North during the ME Period (Fisiak 1996: 90)

	Present indicative		Preterite		Subjunctive (present)		Subjunctive (preterite)	
	Sg	Pl	Sg	Pl	Sg	Pl	Sg	Pl
First	<(e)>		unmarked					
Second	<es> <is>	<es> <is>	<∅>	unmarked	<(e)>	<(en)>	unmarked	unmarked
Third	<es> <is>		unmarked					

Table 16

Inflections of the Imperative, Present Participle, Past Participle and Infinitive of Strong Verbs in the North during the ME Period (Fisiak 1996: 90)

	Imperative		Present participle	Past participle	Infinitive
	Sg	Pl			
First					
Second	unmarked	<es>	and(e)	<(e)n>	unmarked
Third					

The north was characterised by an advanced system of verbal inflections, especially in comparison with that of the south, consisting of, “a highly innovative ... present, with frequent collapse of second- and third-person singular and of both with plural, and an ending in -s for all three collapsed categories” (Lass 1992: 136). Northern features like this one soon spread into other areas (Lass 1992: 134-139), except for some areas in the extreme south (Fischer 1992: 207). However, this system was also conditioned by other factors, collectively known in the literature as the Northern Present Tense Rule (NPTR). Operating mainly in the north, including parts of Scotland, but also in some Midland varieties of ME (Schendl 2000: 264), it stipulated that whether or not the verb took an inflection in the present indicative depended on the type of the subject and its proximity to a lexical verb (Meurman-Solin 1992, Montgomery 1994, King 1997). For example, King (1997: 175; emphasis in the original) demonstrates that the inflections were <(i)s>/<(e)s> in Northern English and <(i)s> in Scots⁵¹ on verbs in the present tense in all persons and numbers, unless “the *subject* [was] *an immediately adjacent personal pronoun* (either preceding or following the verb) which is *first person singular, or first, second or third person plural*”, in which case the verb had no ending. Fulk (2012: 73), too, shows the innovative nature of the system of inflections in the present indicative in the north, more specifically in NME, although with some variation, notably in relation to the second person singular, probably being the result of dialectal variation, “[t]he Northern -es comes eventually to be used as the ending of all persons in the present indicative, including the first person ... [although this inflection] is eliminated ... when the verb is immediately preceded or followed by a personal pronoun, excluding the second person singular.”

Changes in relation to inflectional morphology affected both strong and weak verbs in ME and Scots, but differently (the ME surviving members of OE preterite-present verbs are dealt with below). The grammatical function of strong verbs was formed by having inflectional endings to mark number, subjunctive and participial status, in addition to changes in the root vowel (ablaut) to distinguish between the past and the present (tense). Weak verbs, on the other hand, indicated the distinction between the

⁵¹ King (1997: 175) argues, “in the very earliest Scots, forms in <e-> can occur, but these soon give way to <is> as the regular spelling.”

present and non-present by adding a suffix to a usually invariant root (Fulk 2012: 71). Since the verb **gan** and its variant **can**, as well as the verb **beginnen** belonged to the former Class III of OE strong verbs (*DOE* s.vv. *be-ginnan* v.; *ginnan* v.), an outline of inflectional morphology (inflectional endings and root vowel alternations) of strong verbs in the north, based on the paradigm of the verb **drinken**, is given in Table 17 and Table 18, both tables showing the attrition of inflectional endings on these verbs having taken place.

Table 17

*Inflectional Morphology of **drinken** in the Present Indicative, Preterite, Subjunctive (Present) and Subjunctive (Preterite) in the North during the ME Period (Fisiak 1996: 90-91)*

	Present indicative		Preterite		Subjunctive (present)		Subjunctive (preterite)	
	Sg	Pl	Sg	Pl	Sg	Pl	Sg	Pl
First	drink(e)							
Second	drink-es	drink-es	drank	drank	drink-(e)	drink-(en)	drank	drank
Third	drink-es							

Table 18

*Inflectional Morphology of **drinken** in the Imperative, Present Participle, Past Participle and Infinitive in the North during the ME Period (Fisiak 1996: 90-91)*

	Imperative		Present participle	Past participle	Infinitive
	Sg	Pl			
First					
Second	drink	drink-es	drink-and	drunk-(e)n	drink
Third					

Therefore, in a similar way to new developments occurring in the north in relation to the present indicative, inflectional endings to mark the singular and the plural of both strong and weak verbs were lost quicker in the north than in the south during the ME period. (Fisiak's 1996: 90, 92 list of inflections for strong verbs in the preterite in the north differs from those listed by Fulk's 2012: 74, the discrepancy is probably due to the fact that the former scholar refers to the late ME (IME) period, whereas the latter one to the early ME one. Fulk 2012: 74, for example, demonstrates that the endings <e> and <en> were lost in the indicative of the preterite there after the earliest text, whereas <en> was retained on strong verbs in the past participle). The changes in relation to the root vowel are discussed in detail below.

In OE, strong verbs were divided into seven classes, based on the difference in the number of alternating vowels in the basic forms of the verb: a) the present stem (represented by the infinitive); b) the preterite first and third person singular; c) the preterite second person singular, the plural and the preterite subjunctive; d) the past participle (Fulk 2012: 74). But, the root vowel of strong verbs was affected by levelling in ME. For example, the former Class III of OE strong verbs, to which OE *onginnan* and OE and eME *aginnan*, i.e. the ancestors of ME *gan* and *can*, and OE *beginnan*, the ancestor of ME *beginnan*, belonged, underwent grade reduction in the preterite (Wyld 1914: 205-206, Lass 1992: 130-132, Fisiak 1996: 102, Fulk 2012: 74, 76). This change first occurred in the north, resulting in one vowel being used in the preterite as early as in eME (Fisiak 1996: 102): a) levelling of the preterite singular (second person) under the vowel of the first and third person singular, leading to a more transparent singular/plural opposition; and b) removing the number opposition in the preterite, with the simultaneous generalisation of the singular or the plural vowel (or, even, the vowel of the past participle) (Fisiak 1996: 102) or the extension of the singular stem to the plural (Fulk 2012: 76).⁵² Similar changes took place in Scots. According to (King 1997: 177), the preterite of strong verbs was formed by “modifying the root vowel of the present tense ... In Older Scots, the number of vowel grades per verb had reduced by comparison with Old English, with

⁵² According to Fisiak (1996: 102), the singular and plural dichotomy was preserved in SME, including London, as late as the end of the ME period; there was variation in that both one and two vowels were preserved in the preterite in the Midlands.

the previously different vowels of past tense singular and past tense plural often levelling under the vowel of the singular” (cf. Macafee 1992/93: §8.11).

Although Roberts & Roussou (2003: 43) argue that the loss of infinitival morphology led to the reanalysis of modal (auxiliary) verbs, Warner (2003: 140-144) and Lightfoot (2006: 94) propose that changes concerning the attrition of inflectional morphology of verbs in the present indicative and in the preterite have had consequences on the verbs that were to become modal (auxiliary) verbs, some of the properties also being found typical of emerging non-modal auxiliary verbs at the time (Warner 1993: 103). Firstly, although the OE predecessors of modal (auxiliary) verbs in ME, especially at the beginning of the period, were inflected elsewhere, i.e. *þu cannst, we cunnan, we cuðon*, an exception was the lack of the inflection in the third person singular in the present tense⁵³, an inflectional property of lexical verbs in the present tense (Lightfoot 2006: 94, cf. Warner 1993: 101). As outlined in the preceding paragraphs, inflectional distinctions on verbs, especially in relation to the singular and plural dichotomy, had disappeared during the course of the ME period, firstly in the north. This left the successors of the OE preterite-present verbs isolated; they now looked different from lexical verbs in that they lacked the one morphological feature, namely the inflection in the third person singular, that they shared with the lexical verbs (Lightfoot 2006: 94-95). Secondly, Lightfoot (2006: 95) points out that the morphological distinctiveness of modal (auxiliary) verbs in ME was “a new opacity in their past-tense forms”. This opacity is caused by the fact that the preterite forms of these verbs were phonetically identical to those of the subjunctive in many cases; after the subjunctive forms disappeared, past-tense forms of modal (auxiliary) verbs survived with subjunctive meanings, instead of preterite ones. Lightfoot (2006: 59) concludes, “[s]o *might, could, should*, etc. came to take on new meanings that had nothing to do with past time, residues of the old subjunctive uses; the past-tense forms became semantically opaque.”

⁵³ Some of the ME successors of the OE preterite-present verbs appear with a present indicative plural inflection modelled on lexical verbs. Warner (1993: 101) argues that instead of the historically expected ending for preterite-present verbs, in some parts of the south and south-west Midlands, the <eþ> form, typical of lexical verbs, is found with *can, may* and *shall*, resulting in the following forms *shulleþ, conneþ* and *moueþ*, respectively. The verb *wile*, on the other hand, does not historically belong to the preterite-present group of verbs, with the present tense form *willað* being found in OE.

Therefore, changes in relation to inflectional morphology resulted in modal (auxiliary) verbs looking different from lexical verbs in ME and Scots: a) these verbs became morphologically distinct in that, unlike lexical verbs, they were unmarked in the third person singular in the present indicative (in addition to the attrition of inflectional endings elsewhere, in a similar way to lexical verbs); and b) the preterite forms of these verbs were characterised by semantic opacity, failing to systematically refer to events or actions described as occurring in the past. As an outcome of these changes, this small group of verbs became isolated from others. Lightfoot (2006: 95) maintains that category membership is also based on the morphological and distributional properties of grammatical categories and elements, “[i]f forms look alike and have the same formal, morphological properties, they are usually assigned to the same category; in addition, if forms have the same distribution and occur in the same contexts, they are assigned to the same category.” So, modal (auxiliary) verbs came to be identified as a distinct class on their own. According to Warner (1993: 103), verbs characterised by auxiliation displayed the following morphological characteristics: a) “preterite-present morphology”; b) “the [u]se of past-tense forms without past time reference, outside a motivating context”; and c) “[r]estriction of some of these words to finite forms”. With reference to such characteristics of what were to become modal (auxiliary) verbs, Lightfoot (2006: 95) argues, “the morphological simplification entailed new primary linguistic data and they seem to have triggered new category distinctions. In this case, we know that, following the morphological changes, the surviving verbs of the preterite-present class were assigned to a new grammatical category.”

In summary, the category of the verb was characterised by changes in inflectional morphology during the transition from OE to ME. With reference to the verb *gan* and its variant *can* found in the preterite when in auxiliary use, the relevant changes involved a loss of inflection and modifications to the root vowel in relation to the marking of the category of number in the preterite. However, Lightfoot (2006: 94) argues that changes regarding the morphology associated with the present indicative and the preterite have also had other far-reaching consequences. During this period, a

new class of verbs was also established characterised by auxiliaryhood, i.e. a set of properties setting them apart from the class of lexical verbs. As already outlined in Chapter One, one of the peculiarities of verbs characterised by auxiliation was the restriction of some of these words to finite forms (Warner 1993: 103). This kind of constraint could be observed in the case of *gan* and *can*, but not with *beginnen*, in ME and Scots. Furthermore, this verb and its variant are characterised by reduced inflectional morphology. However, although sources like Denison (1993: 322) and Brinton (1996: 81) argue that *gan* and *can* are morphologically fixed, this claim is made in view that this verb and its variant undergo auxiliation, or even as a kind or subtype of grammaticalization, without references to global language changes affecting the category of the verb in general in ME and Scots, in which *gan* participated, too. In what follows, I present an overview of grammatical forms and inflectional morphology of the verb *gan* and its variant *can*, as well as the verb *beginnen* in ME and Scots, and, subsequently, in the NME and ESc data sets. By doing so, I show that there are striking differences with respect to their morphological paradigms, as a result of decategorialization, analogy and phonological reduction of *gan* and *can*.

3.3 Grammatical Forms and Inflectional Morphology of *Gan*, *Can*, and *Beginnen* in Middle English and Scots

3.3.1 *Gan* and *Can*

3.3.1.1 *Preterite*

When in auxiliary uses, *gan* and *can* only function as finite verbs in ME and Scots, either usually or exclusively, in the preterite (*DSL* s.vv. *can* v. 2; *gan* p.t., *MED* s.vv. *can* v.; *ginnen* v. 3b, *OED* s.vv. *can* v.² I 1a, II; *gin* v.¹, Einkenel 1891: 89, Wuth 1915: 54-57, Taylor 1917, Beschorner 1920, Funke 1922, Häusermann 1930: 18-20, Kerkhof 1966: 154-159, Koziol 1932: 131-134, Mustanoja 1960: 610-615, Smyser 1967: 68, Brinton 1981: 156-158; 1983: 236; 1988: 117-122; 1996: 67-69, Macafee 1992/93: §8.13, Denison 1993: 321-322, J. Smith 1996: 137, Markus 1997, Ogura 1997: 403). However, *can* also has an alternative and additional preterite form *couth*, found “[c]hiefly” in Scots (*OED* s.v. *can* v.²), with Visser (1969: 1579-1580) also having instances of this variant in NME and SME texts. *Couth* is said by the *OED*

(s.v. *can* v.²) to have arisen because *can* as a variant of *gan* was probably seen as a special use of the modal (auxiliary) verb *can*; *couth* was then used instead of *can* to remove the irregularity, whereby *can* in the present indicative in form was used to refer to a past event or situation (cf. Macafee 1992/93: §8.13). However, while the *OED* (s.v. *can* v.²) argues that the variant *couth* “frequently replaced the original *can*” in OSc, Taylor (1917: 577) finds that *couth* was not frequently used in Scots. On the basis of his vocabulary test, which he carries out “to determine what words are frequently repeated as infinitives so as to become conventional phrases or something approaching rime-tags” (Taylor 1917: 576), he speculates whether the infrequency of this variant in Barbour’s *The Bruce* in both the Cambridge and Edinburgh MSS was because “*couth* was just entering the language”, in addition to other options, namely this form being corrupted by the scribes or that this form was being used sparingly by Barbour, appearing more frequently in works by other authors. The spelling variants of *gan* and *can* in ME and Scots, as well as *couth* in Scots⁵⁴, listed here separately for ease of reference, are given in Table 19.

⁵⁴ The *OED* (s.v. *can* v.²) provides spelling variants of *couth* for Scots only, whereas the *MED* (s.v. *can* v.) does not list any forms of *couth* as a variant of *can*.

Table 19

Forms of **gan** and **can** in ME and Scots, as well as **couth** in Scots (DSL s.vv. *can* v. 2a-b; *gan* p.t., MED s.vv. *can* v.; *ginnen* v., OED s.vv. *can* v.²; *gin* v.¹)

	<i>gan</i>			<i>can</i>			<i>couth</i>
	ME		Scots	ME		Scots	Scots
	Sg	Pl	Sg/Pl	Sg	Pl	Sg/Pl	
Preterite	<gain>	<gan>	<gan>	<can>	<can>	<can>	<cought>
	<gan>	<gane>	<gane>	<con>	<con>	<canne>	<could>
	<gane>	<gon>		<conne>	<conne>	<kan>	<couth>
	<gann>	<gonne>		<kan>	<connen>		<cowd>
	<ganne>	<gönnē>			<cun>		<cowld>
	<gannust>	<gonnen>			<cunne>		<cowth>
	<gen>	<gönnen>					<culd>
	<gen>	<gönnenen>					<cuth>
	<gene>	<gun>					<cwth>
	<gon>	<gun>					<qwoth>
	<gond>	<gunde>					
	<göne>	<gunen>					
	<gonne>	<gunne>					
	<gun>	<gunnen>					
	<gune>						
	<gunne>						
	<zan>						

Although both Denison (1993: 322) and Brinton (1996: 81) argue that when in auxiliary use, **gan** is morphologically fixed in the third person singular in the preterite, as part of the process of grammaticalization, this verb and its variant can still be found with inflectional endings in ME, especially in SME (MED s.vv. *can* v.; *ginnen* v., OED s.vv. *can* v.² I 1a; *gin* v.¹; Visser 1969: 1571-1577, Moessner 1989: 37, Sims 2008: 132). For example, Visser (1969: 1576) maintains, “by the side of this *gan/can* there also occurred the plural forms (*cun*), (*gun*), *connen*, *gunne(n)* and *gonne(n)*. They became obsolete long before *gan/can*”. He then provides a number of instances of plural forms of **gan**, but fewer of **can**, from SME, but his only example from NME shows no inflectional endings, as in (1) below. It, therefore, seems that whether **gan** and **can** have inflectional endings varies in the dialect in which this verb and its variant are found. For example, the MED (s.vv. *can* v.; *ginnen* v. 3) lists no examples of inflected **gan** and **can** from NME texts like *Cursor Mundi*, *The Destruction of Troy* and *The Pricke of Conscience*, although it has an instance each

of the former with <e> from *Sir Isumbras* and *The York Plays* (*OED* s.vv. *can* v.² I 1a; *gin* v.¹). This suggests that although singular and plural forms of this verb and its variant occur side by side in SME, invariant *gan* and *can* are likely to be encountered in the northern dialects of this variety. Other sources dealing with *gan* and *can* are not usually concerned with the significance of singular and plural inflectional variation of this verb and its variant in ME, although they do acknowledge, often in passing, the existence of such forms (Mustanoja 1960: 610, Denison 1993: 322, Markus 1997: 349, Sims 2008: 132).

- (1) A *neu liuelade* *cun* *þai* *bigin*
 A new living did they begin
 (13 .. *Cursor Mundi* 2009 (Cotton Vespasian A.3 MS); Visser 1969: 1576)

‘They began a new living’

No references to the preterite singular or plural inflections of *gan* and *can* in auxiliary uses in Scots are usually provided in the literature on the subject (*DSL* s.vv. *can* v. 2; *gan* p.t., Taylor 1917: 573-591, Macafee 1992/93: §8.13, Moessner 1997: 114). However, although this verb and its variant seem to be invariable in form in Scots, some sources list instances of the variant *can* with the final <e>, which might be indicative of the remnant of the plural inflection <en> on verbs in the preterite. Nevertheless, although the *OED* (s.v. *can* v.²) has instances of <canne> beside <can> in this variety, these occur with singular subjects. Both the *DSL* and Macafee (1992/93: §8.13) provide an instance of the same <canne> from *The Scottish Legendary* (taken from *DOST* s.v. *can*), as in (2) below, but this variant is used here with a singular subject used collectively. Therefore, the final <e> appears in these cases to be an instance of orthographic variation. Furthermore, amongst instances of the invariable *gan* and *can* with singular and plural subjects in Scots, Visser (1969: 1576-1577) also has one instance of <gone> with a plural subject from Rolland’s *The Seven Sages*. Nevertheless, although this text is described by Weber (1810: lv-lvi) as “translated by a Scotchman, or at least the language has been considerably altered by some former transcriber of that nation”, it is, crucially, transcribed by an English person, in this respect not being representative of the variety in question, “[t]he scribe, however, who copied the romance into the MS., was probably English; for

though he has retained the peculiarities of the Scottish [sic] dialect, he has considerably anglicized the spelling.” Overall, therefore, the lack of inflections on the verb *gan* and its variant *can* in northern varieties of ME and in Scots appears to be a reflection of the process of attrition of endings on verbs in these varieties.

- (2) A *gret menze com swdanly ... &*
 A *great multitude came suddenly ... &*
 lud canne zele
 loud did/began to yell
 (1380 *The Scottish Legendary* vii 113; *DSL* s.v. *can* v. 2a)

‘A great crowd came suddenly ... and did/began to shout out loudly’

In ME and Scots, the verb *gan* and its variant *can*, as well as *beginnen* are also defined by changes in the root vowel to form other forms, in line with verbs from the former Class III of OE strong verbs. The basic ablaut pattern of vowels of verbs from this class in OE was: a) /e/ (the present stem, best represented by the infinitive); b) /æ/ (the preterite first and third person singular); c) /u/ (the stem of the second person preterite, the plural, and the preterite subjunctive); and d) /o/ (the past or passive participle). However, this class was also characterised by a variety of patterns, with OE *onginnan* and OE and eME *aginnan*, as well as OE *beginnan*, belonging to the *bindan*-type verbs with the following vowel grades: a) /ɪ/; b) /a/; c) /ʊ/; and d) /ʊ/ (Fulk 2012: 78, cf. Fisiak 1996: 102) (these root vowels correspond to those in the basic forms of the verb, as outlined above). In the north, the ablaut pattern of Class III strong verbs in the preterite was initially /a/ or /a:/ in the first and the third person singular, and /ʊ/ and /u:/ in the second person singular, the plural and the subjunctive before nasal consonants (Fulk 2012: 78). As already stated in Section 3.2, a number of changes took place in the north, affecting the root of Class III of OE strong verbs. The *OED* (s.v. *begin* v.¹) proposes this more specifically in relation to *beginnen*, “there was an early tendency to level the forms of the 1st and 3rd singular *began*, and of the 2nd singular *begunne*, plural *begunnon*, which has resulted in the establishment of *began* as the standard form; but, an alternative from the old plural *begun* has also come down to the present day”. Therefore, if levelling was not completed before the plural forms of *gan* and *can* disappear in the north, then there would be variation between the singular and plural forms of this verb and its variant,

just as it would be expected to have this kind of variation for *beginnen* before levelling was completed.

The spelling variants representing corresponding root vowels found with *gan* and *can* in ME and Scots, as well as *couth* in Scots, listed here separately for ease of reference, are provided in Table 20.

Table 20

Spelling Variants of the Root Vowel of gan and can in ME and Scots, as well as couth in Scots (DSL s.vv. can v. 2a-b; gan p.t., MED s.vv. can v.; ginnen v., OED s.vv. can v.²; gin v.¹)

	<i>gan</i>			<i>can</i>		<i>couth</i>	
	ME		Scots	ME	Scots	Scots	
	Sg	Pl	Sg/Pl	Sg	Pl	Sg/Pl	
Preterite	<a>					<ou>	
	<ai>	<a>			<a>	<ow>	
	<e>	<o>	<a>	<a>	<o>	<a>	<u>
	<o> ⁵⁵	<u>			<u>		<w>
	<u>						<wo>

It can be seen from this table that the number of vowel grades per *gan* and *can*, as represented by spelling variants, is reduced in Scots, while the wider range of them is still found with this verb and its variant in ME (spelling variants corresponding to the root vowel in *couth* will be dealt with below). Furthermore, the data in the table shows that spelling variants corresponding to vowels used to mark the singular and plural dichotomy on *gan* and *can* can still be found in ME, whereas this is not the case in Scots. However, although the spelling variants of *gan* and *can* can indicate

⁵⁵ The <o> forms in the preterite singular in the table are not mentioned in the *OED*. This could be accidental, though: a) if both <a> and <u> orthographic variants are attested, this does not appear to be like a formal distinction, although this depends on how this verb and its variant distribute – same author/variety or different authors/varieties; and b) the *MED* (s.v. *ginnen* v.) also lists <o> and <e> for the singular.

the likely root vowel, the sources used to compile this list, namely the *DSL* (s.vv. *can* v. 2a-b; *gan* p.t.), the *MED* (s.vv. *can* v.; *ginnen* v.) and the *OED* (s.vv. *can* v.²; *gin* v.¹), do not differentiate between forms of this verb and its variant in auxiliary uses, on the one hand, and *ginnen* as a lexical verb, on the other. Furthermore, the *MED* (s.v. *ginnen* v. 3b) argues that *ginnen* in the present indicative found with the plain infinitive can also have auxiliary uses to denote events or situations as occurring, rather than focussing on the inception of events or situations. Nevertheless, since *gan* and *can* mainly occur with the preterite meaning when they are in auxiliary use, as has been argued by most sources dealing with this verb and its variant, on the basis of this, it can be concluded that spelling variants denoting root vowels, for example, in past participle forms of *gan* (for example, <o> in *gonnen*, <u> in *gunen* and <i> in *ginnen*, as listed in the *MED*) are likely to belong to *ginnen* in ME (the *DSL* s.v. *gan* p.t. mentions no present indicative forms of *ginnen* in Scots). Finally, this problem of potential confusion whether spelling variants illustrating vowel grades represent *gan* in auxiliary use and *ginnen* as a lexical does not appear to be relevant to *can*; this is because no non-finite forms of this variant have been found to occur in ME and Scots.

In comparison to the above findings from the literature on the subject, the results of this investigation, as presented in Table 21, Table 22 and Table 23, show that *gan* and *can* only occur in the preterite in the NME and ESc data sets, as in (3) below, (4) below and (5) below.

Table 21⁵⁶

Forms of gan with Infinitive in the NME Data Set in Period I and Period II

		<i>gan</i>					
		Sg			Pl		
Person	Form	<i>N</i>	%	Form	<i>N</i>	%	
Period I	First	-	-	-	-	-	
	Second	-	-	-	-	-	
	Third	<gan> <gun>	11 2	65 12	<gan>	4	24
Period II	First	-	-	-	-	-	
	Second	<gan>	1	1	-	-	
	Third	<gan> <gane>	58 1	77 1	<gan> <gane>	13 2	17 3

Notes: *N* – Observed absolute frequency
% – Per-cent of occurrence of *gan* per person per number per period

Table 22

Forms of can with Infinitive in the NME Data Set in Period I and Period II

		<i>can</i>					
		Sg			Pl		
Person	Form	<i>N</i>	%	Form	<i>N</i>	%	
Period I	First	<can>	1	13	-	-	
	Second	-	-	-	-	-	
	Third	<can>	6	75	<can>	1	13
Period II	First	-	-	-	-	-	
	Second	-	-	-	-	-	
	Third	<can>	8	67	<can>	4	33

⁵⁶ Instances of *gan* and *can*, as well as *beginnen* used with collective nouns are subsumed under the heading ‘plural’, where relevant.

Notes: *N* – Observed absolute frequency
 % – Per-cent of occurrence of **can** per person per number per period

Table 23

Forms of can with Infinitive in the ESc Data Set in Period I

		<i>can</i>					
		Sg			Pl		
Period I	Person	Form	<i>N</i>	%	Form	<i>N</i>	%
	First	<cane>	1	13	-	-	-
	Second	-	-	-	-	-	-
	Third	<can> <cane>	2 3	25 38	<can>	2	25

Notes: *N* – Observed absolute frequency
 % – Per-cent of occurrence of **can** per person per number per period

- (3) *And say “weles us þat here er wonnand”,
 And say “choose us that here are living”,
 And thank God omang þat þam gun wysse
 And thank God together that them did point
 Til mekenes, þat þam led til þat blysse
 To meekness, that them led to that bliss
 (1350-1420 *The Pricke of Conscience* p. 250; *HCET*)*

‘and said, “choose us that are living here”, and together thanked God, who pointed them to meekness, and led them to that bliss’

- (4) *Dis herd dauid, and forth can stand,
 This heard David, and forth did stand,
 “Sir king,” he said, “hald me couenand,
 “Sir king,” he said, “hold me covenant,
 I tru truli in godds might
 I trust truly in God’s might
 (1350-1420 *Cursor Mundi* p. 432; *HCET*)*

‘David heard this, and stepped forward, “Sir king,” he said, “keep promise to me, I truly have faith in God’s might’

- (5) *Sad to hyme agane, “Thu ga
 Said to him again, “Thou go
 Til hyre that cane this demand ma
 To hir that did this demand make*

And spere at hyre grathly.
And ask at her promptly.
 (1375-1420 *St. Andrew and the Three Questions* l. 1084-1087;
TEAMS)

‘[the porter] said to him again, “Go to her that made this demand and ask her immediately.”’

In terms of forms of *gan* and *can*, the following observations can be made in relation to this verb and its variant in the NME data set: a) *gan* is mainly realised as the invariable <gan> in Period I (89%) and Period II (95%), including instances of this verb realised as <gun> (without an inflectional suffix, too) in Period I, too, as in (6) below (12%). However, in Period II, in addition to instances of the invariable <gan> (95%), there are also examples of <gane> (4%), i.e. with the final <e>, as in (7) below; b) *can* exclusively occurs as the invariable <can> in Period I and Period II, as in (4) above. With regards to *can* in the ESc data set, this variant is found as either <can> (50%), as in (8) below, or <cane> (50%), as in (5) above, in the ESc data set. Although the dropping of <e> first began in the north in the 12th century (Fisiak 1996: 86), instances of *gan* in the NME data set, as in (7) below, and *can* in the ESc data set, as in (5) above, are still used with <e>, even in Period II; however, this variation appears to be orthographic. For example, in the case of <gane>, the subject is in the third person plural, as in (7) below, whereas in the case of <cane>, it is in the third person singular, as in (5) above. Overall, *gan* and *can* take no inflectional suffix in the NME and ESc data sets, otherwise typical, to an extent, of the plural forms of the former Class III of OE strong verbs of the *bindan*-type in eME. Instead, this verb and its variant show some orthographic variation in both data sets. On the basis of this, the attrition of inflectional morphology associated with number marking is complete on this verb and its variant in the texts included in the NME and ESc data sets. The lack of separate inflections to mark the singular and the plural is, therefore, in line with my expectations, following descriptions on the lack of inflectional suffixes of this verb and its variant in the north at the start of the period.

(6) *And sun on þam sa gun it proue,*
 And soon on them so did it prove,
 For was þar þan na langer bide
 For was there then no longer delaying

(1350-1420 *Cursor Mundi* p. 424; *HCET*)

‘And it soon demonstrated to them, for then there was no longer delaying’

- (7) *And the walles downe gan he dyngē.*
And the walls down did/began to he strike.
And than gane alle the pepille crye.
And then did/began to all the people call.
Unto God and to mylde Marye
Unto God and to mild Mary
With sorowe and grete wepyngē.
With sorrow and great weeping.
(1420-1500 *Octavian* l. 883-886; *TEAMS*)

‘And he did/began to break the walls down and then all the people did/began to call to God and to merciful Mary with sorrow and great weeping.’

- (8) *Sancte Julyane than and his wyf*
Saint Julian then and his wife
To God led thai sa thankful lyfe –
To God led they so thankful life –
That thare dedis sa wele can stere
That their deeds so well did guide
The lytil tyme that tha lifit here –
The little time that they lived here –
(1375-1420 *The Life of St. Julian Hospitaller* l. 474-476; *TEAMS*)

‘Saint Julian and his wife then led a life pleasing to God, who guided their actions well during the short time they lived here’

The findings of this study also demonstrate that in the NME and ESc data sets, **gan** and **can** almost exclusively occur with the orthographic variant <a>, representing the root vowel /a/, as in (9) below and (10) below. The occurrence of this verb and its variant with this type of variant corresponding to this vowel is typical of the process of levelling of the preterite singular of all such verbs in the north under the vowel of the first and third person singular, with the number opposition being removed at a later date (Wyld 1914: 205-206, Lass 1992: 130-132, Fulk 2012: 74, 76). However, as shown in the preceding paragraph, there is also some minor variation in the root vowel of **gan** in the NME data set in Period I only, as exemplified by the occurrence of this verb with <u>, as in (3) above and (6) above. Nevertheless, this seems to be orthographic, too, rather than being the remnant of the OE preterite plural vowel

alternation: a) these two instances of <gun> are used with third person singular subjects here; and b) one of the instances of <gun> occurs alongside <gan> in the same text (although not in a nearby part of the text), namely *Cursor Mundi*, as in (6) above and (11) below.

- (9) *Adam, thyselffe* *made al þis syte,*
 Adam, thyself made all this grief,
For to the tree þou wente full tyte
 For to the tree thou went full quickly
And boldelyon the frute gan byte
 And boldly on the fruit did/began to bite
My lord forbed
 My lord forbade
 (1420-1500 *The York Plays* p. 70; *HCET*)

‘Adam, you made all this grief yourself, because you very quickly went to the tree and boldly did/began to bite on the fruit that my lord forbade’

- (10) *þat it was iewus þan wist i wel,*
 That it was Jesus then knew I well,
And cuthli for him can i knele,
 And certainly for him did/began to I kneel,
And i said, ‘ai lauerd! blisced þou be
 And I said, ‘Aye Lord! Blessed you be
 (1350-1420 *Cursor Mundi* p. 1014; *HCET*)

‘I knew well that it was Jesus, and certainly I did/began to genuflect/kneel before him, saying, ‘Yes, Lord! Blessed be you’

- (11) *þe slogth he fand þat him gan wiss*
 The path he found that him did mark
Tilward þe zate of paradis;
 Toward the gate of paradise;
 (1350-1420 *Cursor Mundi* p. 82; *HCET*)

‘He found the path that marked the way for him towards the gate of paradise;’

Furthermore, as already mentioned, additional developments in relation to the root vowel are recorded for *can* as a variant of *gan*, whereby “[i]n Older Scots ... *couth, cowd, could* [as preterite forms of the modal (auxiliary) verb *can*] ... frequently replaced [the variant *can*] ... in its function as a periphrastic auxiliary of the past tense, in order to remove an apparent irregularity where what was perceived

as a present stem of *can* [as a modal (auxiliary) verb] was used as a past tense in the periphrastic construction” (*OED* s.v. *can* v.², cf. *DSL* s.v. *can* v. 2). Therefore, both *can* and *couth* are forms associated with expression of the preterite, but with some variation in the root vowel, the former having the spelling variant <a> corresponding to the root vowel /a/, whereas the latter displays a number of spelling variants denoting /ʊ/. However, as already stated in Chapter Two, no instances of *couth* are found in the NME and ESc data sets.

3.3.2 *Beginnen*

Unlike the verb *gan* and its variant *can*, the verb *beginnen* occurs in a range of finite and non-finite forms throughout ME and Scots, typical of lexical verbs at the time (*DSL* s.v. *begin* v., *MED* s.v. *biginnen* v., *OED* s.v. *begin* v.¹, Visser 1969: 1373-1375, 1580-1581, Brinton 1981: 181-186; 1983; 1988: 158-159, Macafee 1992/93: §8.11, Ogura 1997, Sims 2008: 151-162; 2014). More specifically, this verb is found in descriptions in the literature as functioning in the present indicative, preterite, imperative, subjunctive, infinitive, as in (12) below, but also as the present participle and the past participle, as in (13) below (*DSL* s.v. *begin* v., *MED* s.v. *biginnen* v., *OED* s.v. *begin* v.¹, Fulk 2012: 72-81).

- (12) *Here is best now to begyn Thi purpos*
Here is best now to begin thy purpose
 (c1420 Andrew of Wyntoun *The Original Chronicle of Scotland* VII
 63; *DSL* s.v. *begin* v. 2a)

‘Here is best to begin your purpose now’

- (13) *Bringe to ende that thou hast by-gonne*
Bring to end that thou has begun
 (1307 *Elegy on the Death of Edward I* viii; *OED* s.v. *begin* v.¹ 2a)

‘Bring to an end what you have begun’

As expected of verbs at the time, *beginnen* also displays dialectal variation with reference to inflectional endings, not only on the basis of the differences between the north and the south, but also within the Northumbrian *Sprachgebiet*, i.e. between NME and Scots. Given that this verb belongs to the former Class III of OE strong

verbs, too (*DOE* s.vv. *be-ginnan* v.; *ginnan* v.), the system of inflections on ***beginnen*** in ME and Scots to mark finite forms (present indicative, preterite, imperative) and non-finite ones (present participle, past participle and infinitive) is essentially the same as that of other strong verbs in these respective dialects. A list of finite forms of this verb in ME and Scots, including ***begouth*** in Scots, listed here separately for ease of reference, is provided in Table 24, whereas that of non-finite ones is given in Table 25 (the sources used to compile these forms do not distinguish between northern dialects of ME and southern ones in respect to ***beginnen*** or provide separate forms of this verb in the imperative). The following emerge from these tables: a) distinctions into the singular and plural of this verb in the preterite are still present in ME, but not in Scots; and b) the preterite of this verb displays variation in the root vowel in ME and Scots, although differently in both varieties, as discussed below.

Table 24⁵⁷

Finite Forms of *beginnen* in ME and Scots, and *begouth* in Scots (DSL s.v. *begin* v., MED s.v. *biginnen* v., OED s.v. *begin* v.¹)

	<i>beginnen</i>		<i>begouth</i>	
	ME		Scots	Scots
	Sg	Pl	Sg/Pl	Sg/Pl
Present Indicative	<begin>		<beggine>	
	<bigine>		<begine>	
	<biginne>		<begyne>	
	<biginnen>		<biggin>	
	<bigunen>			
	<bigyn>			
	<bigyne>			
	<bygyn>			
	<bygynne>			
	Preterite	<bigan>	<began>	<began>
<bigane>		<begann>	<begane>	<becuth>
<bigon>		<beganne>	<beggan>	<begod>
<bigön>		<begunnen>	<bygane>	<begoud>
<bygan>		<bigan>		<begoude>
<bygon>		<biganne>		<begould>
<bygonne>		<bigönne>		<begouth>
		<bigönnen>		<begouthe>
		<bigun>		<begoutht>
		<bigunne>		<begovd>
		<bigunnen>		<begovid>
		<bygonne>		<begowth>
		<bygun>		<begowthe>
		<bygunne>		<begud>
				<begude>
				<beguid>
				<beguith>
				<beguld>
				<beguth>
				<begutht>
			<begwd>	
			<begwid>	

⁵⁷ In modern editions of MSS, compound verbs like OE *onginnan* and OE *beginnan* are sometimes hyphenated, but usually written as two separate words in MSS (OED s.v. *on-* prefix 1b). In the corpora used in this study, the practice of hyphenating is still sometimes present, but is not uniformly applied to all forms of *beginnen* in the NME and ESc data sets. I have, therefore, removed diacritic marks separating the prefix *be* from the root *ginnen* from the tables to ensure consistency.

Table 25

*Non-finite Forms of **beginnen** in ME and Scots (DSL s.v. *begin* v., MED s.v. *biginnen* v., OED s.v. *begin* v.¹)*

		<i>beginnen</i>		
		ME	Scots	
Present participle		<beginnand(e)>		
		<begonnen>	<beginning>	
		<begun>	<beginnyng>	
		<begunnen>	<begon>	
		<bigan>	<begone>	
		<bigönne>	<begonne>	
		<bigönnen>	<begonnyn>	
		<bigun>	<begoud>	
		<bigunne>	<begoun>	
		<bigunnen>	<begoune>	
	Past participle		<bygonne>	<begown>
			<bygun>	<begownne>
			<bygunne>	<begun>
				<begune>
			<begunnin>	
			<begunyn>	
			<begunyng>	
			<begwin>	
			<begwn>	
			<begwne>	
		<bygonnyne>		
Infinitive	<begin(n)(e)>			

In contrast to *gan* and *can*, descriptions in the literature show that *beginnen* occurs with a wider range of spelling variants, depending on the grammatical function, denoting respective root vowels in both ME and Scots. Furthermore, the choice of spellings associated with the root vowel also shows that the singular and plural dichotomy in relation to this verb in the preterite is retained in ME, but not in Scots. The list of spelling variants corresponding to the root vowel of *beginnen* in ME and Scots, including *begouth* in Scots, listed here separately for ease of reference, is provided in Table 26.

Table 26

*Spelling Variants of the Root Vowel of **beginnen** in ME and Scots, and **begouth** in Scots (DSL s.v. *begin* v., MED s.v. *biginnen* v., OED s.v. *begin* v.¹)*

	<i>beginnen</i>		<i>begouth</i>	
	ME		Scots	Scots
	Sg	Pl	Sg/Pl	Sg/Pl
Present indicative	<i>		<i>	
	<u>		<y>	
	<y>			
Preterite	<a>	<a>	<a>	<ou>
	<o>	<o>	<o>	<ow>
		<u>		<u>
				<ui>
				<w>
Past Participle	<a>		<i>	
	<o>		<o>	
	<u>		<ou>	
			<ow>	
			<u> <w>	

The following specific observations can be made in relation to the data from this table: a) in relation to ME, the choice of spelling variants representing the root vowel of *beginnen* in the preterite appears to be similar to that of *gan* and *can* in this variety. Although <a>, <o> and <u> found to denote the root vowel are shared by these verbs in this variety, these verbs significantly differ in the choice of spellings of the vowel grade in the singular: while *gan* shows a greater degree of variation in this respect, both *can* and *beginnen* display the opposite tendency; and b) in relation to Scots, the choice of spellings denoting the root vowel of *beginnen* in Scots appears to be wider, with the variant <o> being found in addition to the variant <a>, which is the only one used to denote the root vowel in *gan* and *can* in auxiliary uses in this variety. Overall, therefore, while the choice of spellings denoting the root vowel on *beginnen* is generally more varied in ME and Scots, in comparison to *gan* and *can*,

this is mainly due to the fact this verb is used in both finite and non-finite contexts in these varieties. In relation to the preterite, the choice of vowel grades on *beginnen*, as denoted by spelling variants, varies with respect to the dialect of ME, although in Scots, this verb shows a tendency for more variation than *gan* and *can* do.

With reference to the NME and ESc data sets, the verb *beginnen*, formed by prothesis of the prefix *be* (forming derivative verbs with the sense of ‘about’ and ‘around’) to the root *ginnen* (*OED* s.v. *begin* v.¹), is characterised by occurrence in finite and non-finite forms, although to a varying degree: the present indicative and preterite, subjunctive and imperative, as well as the present and past participles, and the infinitive. As already mentioned, this verb also seems to be associated in form with *couth*, as the preterite form of *can* as a variant of *gan*. For example, the *OED* (s.v. *begin* v.¹) states, “[t]he Scots forms *begouth*, *begoud*, seem due to some form-association with *couth*, *could*, probably through the aphetic form *gan*, which became in Scots *can*, and was thus identical in form with *can* ‘to be able’” (*DSL* s.v. *begin* v.). Both the *OED* and the *DSL* then provide some examples of *begouth* from ESc, as in (14) below and (15) below. However, as will become clear in Chapter Five, this variant differs from *couth* in terms of meaning, too.

- (14) *Thair hertis all begouth to fale*
 Their hearts all began to fail
 (a1380 (1487) John Barbour *The Bruce* IX 183; *OED* s.v. *begin* v.¹ 1a)

‘All their hearts began to fail’

- (15) *Thay begould to requyre that Messe sould be*
 They began to require that Mass should be
sett up agane
 set up again
 (a1572 Knox *The Works of John Knox: The History of Reformation in Scotland* i 389; *OED* s.v. *begin* v.¹ 1a)

‘They began to demand that Mass should be set up again’

Overall, descriptions in the literature suggest that *beginnen* still maintains a distinction between the singular and plural vowels in the preterite in ME, but not in Scots. Further developments in terms of the root vowel of this verb can be observed

in Scots only, believed to be on analogy to *couth* as the preterite form of *can* as a variant of *gan*. In what follows, I present an overview of all the grammatical forms of *beginnen* in which this verb is found in the NME and ESc data sets.

3.3.2.1 Present Indicative and Preterite

The types of inflectional suffixes found with the verb *beginnen* in the present indicative in the NME and ESc data sets are provided in Table 27 and Table 28, respectively.

Table 27

*Inflections of **beginnen** in the Present Indicative in the NME Data Set in Period I and Period II*

		<i>beginnen</i>					
		Sg			Pl		
Person	Suffix	<i>N</i>	%	Suffix	<i>N</i>	%	
Period I	First	<∅>	1	5	<∅>	2	9
	Second	-	-	-	-	-	-
	Third	<es>	11	50	<∅>	4	18
		<s>	2	9	<s>	1	5
Period II	First	-	-	-	-	-	
	Second	-	-	-	-	-	
	Third	<es>	6	55	<∅>	1	9
		<ys>	4	36			

Notes: *N* – Observed absolute frequency
 % – Per-cent of occurrence of inflection of *beginnen* per person per number per period

Table 28

*Inflections of **beginnen** in the Present Indicative in the ESc Data Set in Period I and Period II*

		<i>beginnen</i>						
		Sg			Pl			
Period I	Person	Suffix	N	%	Suffix	N	%	
		First	-	-	-	-	-	-
		Second	-	-	-	-	-	-
		Third	-	-	-	-	-	-
Period II		First	-	-	-	-	-	
		Second	-	-	-	-	-	
		Third	<is> <ys>	10 2	83 17	<∅> <is>	1 5	17 83

Notes: N – Observed absolute frequency
% – Per-cent of occurrence of inflection of *beginnen* per person per number per period

In the present indicative, inflectional endings on *beginnen* range from <∅>, <(e)s> to <ys> in the NME data set, as in (16) below, and from <∅>, <is> to <ys> in the ESc data set, as in (17) below. The choice of inflectional endings on *beginnen* in the these data sets, therefore, appears to be in line with those provided in the literature, where Fisiak (1996: 90) and King (1997: 175) have <(i)s>/<(e)s> for verbs in all persons and numbers in Northern English, whereas King (1997: 175) identifies <(i)s> in OSc (cf. Macafee 1992/93: §8.2). Furthermore, whether or not this verb takes inflectional endings also depends on the working of the NPTR, which operated on all lexical verbs in NME and in Scots, as well as some Midland varieties of ME (Schendl 2000: 264, cf. Meurman-Solin 1992, Montgomery 1994, King 1997). For example, *beginnen* takes <(e)s> with non-pronominal plural subjects in the NME data set in Period I, as in (18) below, and <is> with the same type of subject in the ESc data set in Period II, as in (19) below. On the other hand, in accordance with the NPTR, this verb has no inflection with adjacent pronominal plural subjects in the NME data set

in Period I, as in (20) below, and in Period II, as well as in the ESc data set in Period II. Overall, the findings on the verb *beginnen* in the present indicative in the NME and ESc data sets show that this verb behaves like other lexical verbs at the time: it has inflectional endings (variation in the orthographic form of suffixes is variety-specific), depending on the person and number, but also in accordance with the NPTR, i.e. the type and proximity of the subject in relation to the verb.

- (16) *Þe first: is writen bifore, & bigynnes at*
 The first: is written before, & begins at
Ante oracionem prepara a[nimam] t[uam] &
 Ante oracionem prepara a[nimam] t[uam] &
lastis h[id]er
 lasts hider
 (1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 142;
ICoMEP)

‘The first is written before, and begins at *Ante oracionem prepara animam tuam* [before prayer prepare thy soul] and continues here’

- (17) *HER BEGYnNYS A NOBYL TRETYSE MADE*
 Here begins a noble treatise made
OF A GUD PHEsICIAN JOHN OF BURDOUSE
 of a good physician John of Burdouse
FOR MEDICENE AGAYNE YE PESTILENS jWYLL
 for medicine against the pestilence evil
 (1420 *Liber S. Marie de Calchou* 164; *LAOS*)

‘Here begins a great book written by a good physician John of Burdouse on medicine against the evil’

- (18) *Þen many begynnes þe thyng þat þai may*
 Then many begin[s] the thing that they may
neuer-mare bryng till ende
 never-more bring to end
 (1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 6;
ICoMEP)

‘Then many begin the undertaking that they may no longer bring to an end’

- (19) *And than begynnys drouthis and grete hete of*
 And then begins droughts and great heat of
sonne to regne – and wyndis softis and temperis
 son to reign – and winds softs and tempers
 (1420-1500 *The Prose Works of Sir Gilbert Hay* p. 99; *HCOS*)

‘And then droughts begin and great heat radiates from the sun – and winds become soft and moderate’

- (20) *Dis gylder layes oure enmy to take vs with,*
This trap lays our enemy to take us with,
when we begyn to hate wyckednes
when we begin to hate wickedness
(1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 6;
ICoMEP)

‘Our enemy lays this trap to catch us, when we begin to hate wickedness’

The verb ***beginnen*** has no inflectional suffixes to distinguish the singular from the plural in the preterite in the NME and ESc data sets, as presented in Table 29 and Table 30, respectively.

Table 29

*Forms of **beginnen** in the Preterite in the NME Data Set in Period I and Period II*

		<i>beginnen</i>					
		Sg			Pl		
Person	Form	<i>N</i>	%	Form	<i>N</i>	%	
Period I	First	<bygane>	2	18	-	-	-
	Second	-	-	-	-	-	-
	Third	<began> <bigan>	2 7	18 64	<bigan>	2	100
Period II	First	<began>	3	3	-	-	-
		<begane>	2	2	-	-	-
	Second	-	-	-	-	-	-
	Third	<becan>	1	1			
		<began>	96	81	<began>	16	67
		<begane>	5	4	<begane>	4	17
		<begon>	1	1			
		<bigan>	1	1			
		<bygan>	7	6	<bygan>	3	13
		<bygane>	2	2			
<bygann>		1	1				
				<hegan>	1	4	

Notes: *N* – Observed absolute frequency
 % – Per-cent of occurrence of **beginnen** per person per number per period

Table 30

*Forms of **beginnen** in the Preterite in the ESc Data Set in Period I and Period II*

		<i>beginnen</i>					
		Sg			Pl		
	Person	Form	<i>N</i>	%	Form	<i>N</i>	%
Period I	First	-	-	-	<began>	1	50
		-	-	-	<begouth>	1	50
	Second	-	-	-	-	-	-
	Third	-	-	-	-	-	-
Period II	First	-	-	-	<begane>	1	25
		-	-	-	-	-	-
	Third	<beggan>	1	25	<begane>	1	25
		<begouthe>	1	25	<begouth>	2	50
		<begwd>	1	25			
		<bigane>	1	25			

Notes: *N* – Observed absolute frequency

% – Per-cent of occurrence of **beginnen** per person per number per period

The lack of inflectional ending <e> and <en> to signal the plural of **beginnen** in the preterite is to be expected, given statements in the literature about the loss of this kind of inflectional marking on verbs in this function in the north “after the earliest texts” (Wyld 1914: 205-206, Lass 1992: 130-132, Macafee 1992/93: §8.11, Fisiak 1996: 102, King 1997: 177, Fulk 2012: 74). However, some instances of this verb appear to occur with the final <e> in the NME and ESc data sets, as in (21) below and (22) below, respectively. This might suggest that this would be a remnant of the <en> inflection for the plural, even if it is not consistently applied throughout the whole paradigm. The final <e> is, therefore, found on **beginnen** regardless of the person and number of the subject, as in (23) below and (24) below. In addition, the final <e> occurs on **beginnen** alongside instances of this verb without it in the following texts, suggesting that its use there is orthographic: *The Mirror of St. Edmund, Octavian, Richard Rolle of Hampole, Yorkshire Writers* and *York Plays* in

the NME data set and De Irlandia's *The Meroure of Wyssdome* in the ESc data set (sometimes even in the same line, but with different subjects). This does not surprise, as the final <e> was being lost in the unstressed words from the 11th century onwards and disappeared in stressed words in the 13th century in the northern varieties of ME (J. Smith 1996: Intralinguistic Correspondences I: The Loss of Final -E), whereas in Scots in the 14th century (Duffell 2008: 96). Overall, *beginnen* in the preterite occurs with no inflectional endings to mark the distinction between the singular and the plural in the NME and ESc data sets.

- (21) *þey lokede to oure lady þat scho sulde dyghte þe*
they looked to our lady that she should wipe the
heuede, and þane begane they to wepe all
head, and then began they to weep all
newe.
newly.
(1420-1500 *Richard Rolle of Hampole, Yorkshire Writers* p. 210;
ICoMEP)

‘they turned to our lady so that she would wipe the head [of Jesus], then they all began to weep immediately.’

- (22) *TO YE PLACE QUHAR YAI BEGANE*
to the place where they began
QUHAR IT WES jNQUIRIT BE YE SAID
where it was inquired by the said
jUGE
judge
(1446 *Registrum Episcopatus Aberdonensis* 322; *LAOS*)

‘to the place where they began, where it was questioned by the said judge’

- (23) *And by þat tyme it smote one after*
And by that time it smote one after
mydnyghte – and whene I bygane to slepe it
midnight – and when I began to sleep it
smote elleuene
smote eleven
(1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 386;
ICoMEP)

‘And by that time it struck one after midnight – and when I began to sleep it struck eleven’

- (24) *And quhen wys~dome begouthe in grece, as*
And when wisdom began in Greece, as
jn the tyme of tebez, jasone and hercules,
in the time of Tebez, Jasone and Hercules,
þan begouth þar nobile cheualry and
then began there noble chivalry and
worthines
worthiness
 (1420-1500 De Irelandia *The Meroure of Wyssdome* p. 9; *HCOS*)

‘And when wisdom began in Greece, as in the time of Tebez, Jasone and Hercules, then noble chivalry and worthiness began there’

In relation to the root vowel, the findings of this study show that the preterite of the verb *beginnen* is formed with <a> exclusively in the NME data set, as in Table 29, and predominantly in the ESc data set, as in Table 30. The vowel alternation in the preterite is, therefore, under the levelling of the vowel found on this verb in the first and third person singular throughout the rest of the paradigm of this verb, eventually leading to the removal of number opposition, as described in the literature on the simplification of the complexity of vowel alternations taking place in the north at the beginning of the ME period. Furthermore, the developments in relation to the root vowel of *beginnen* found in Scots only, as already described in Section 3.3.2, can also be observed on this verb in the ESc data set. In addition to <a>, the choice of orthographic variants of *beginnen* in the preterite also includes <ou>⁵⁸/<w>, as in (25) below and (26) below, corresponding to the root vowel /u:/⁵⁹. Instances of *begouth* appear to be formed on the analogy of the preterite *couth* (not found in the data sets used here, but reported by the *OED* s.v. *can* v.² to be chiefly found in Scots, with Visser 1969: 1579-1580 having instances of this variant in NME and SME texts, too), namely the preterite form of *can* as a variant of *gan*, rather than indicative of the presence of the plural form of *beginnen*, i.e. /u/, in addition to /a/ and /o/. Overall, therefore, the findings show that *beginnen* in the preterite is formed by changes in the root vowel in the NME and ESc data sets. While this root vowel is

⁵⁸ In addition, the other French variant <ow> was also found to be adopted by the 14th century, but is not attested in instances of *beginnen* in the NME and ESc data sets.

⁵⁹ Although <begouth(e)> is found with the plural subject *we* in the *LAOS* in Period I, as in (25), and with both singular *wys~dome* ‘wisdom’ and conjoined *nobile cheualry and worthines* ‘noble chivalry and worthiness’ subjects in the *HCOS*, as in (24), whereas <begwd> with the singular subject *YIS LARD OF MELDRUM* ‘this Lord of Meldrum’ in the *LAOS* in Period II, these forms appear to be orthographic variants of *begouth*.

represented by the orthographic <a> in these data sets, there is also some variation in the ESc data set, with <ou> and <w> being used in some cases, too.

- (25) *YT BUR[NE] TIL IT ENTer IN YE BURNE*
 that burn to it enter in the burn
OF BRYNnes AND SWA TIL IT ENTer IN
 of Brynnes and so to it enter in
GRIGis FURDE QUHAR WE BEGOUTH
 Grig’s ford where we began
FIRST
 first
 (1400 *Registrum Episcopatus Aberdonensis* 179; LAOS)

‘that burn enters Bishop Brynnes’ burn, which then enters Grig’s ford, where we started’

- (26) *AN-OTHur LARD & LARGLY BEGWD QUHAR*
 another lord & largely began where
HIS PREDECESSOR LEFTE ERYT &
 his predecessor left ploughed &
LABORIT YE SAID LAND
 laboured the said land
 (1459 *Arbroath Abbey, “Registrum Nigrum”* 341; LAOS)

‘another lord and largely began where his predecessor left ploughing and labouring the said land’

3.3.2.2 *Subjunctive and Imperative*

The verb *beginnen* is also found in other finite forms: the subjunctive and imperative as an expression of non-fact (mood) in the NME and ESc data sets, as in Table 31 and Table 32, respectively; the number of instances of this verb in these functions is considerably lower than in the present indicative and in the preterite.

Table 31

*Forms of **beginnen** in the Subjunctive and Imperative in the NME Data Set in Period I and Period II*

		beginnen								
		Subjunctive			Imperative					
		Sg			Sg			Pl		
		Form	<i>N</i>	%	Form	<i>N</i>	%	Form	<i>N</i>	%
Period I	<bigyn>	1	100				<begynne>	1	25	
				<bigyn>	2	50	<bigyn>	1	25	
Period II	-	-	-				<begynn>	1	33	
				<begynne>	1	33	<begynne>	1	33	

Notes: *N* – Observed absolute frequency
% – Per-cent of occurrence of **beginnen** per number per period

Table 32

*Forms of **beginnen** in the Subjunctive in the ESc Data Set in Period I and Period II*

		beginnen		
		Subjunctive		
		Pl		
		Form	<i>N</i>	%
Period I	-	-	-	
Period II	<begyn>	1	50	
	<begyne>	1	50	

Notes: *N* – Observed absolute frequency
% – Per-cent of occurrence of **beginnen** per number per period

When used in the subjunctive, **beginnen** is not inflected, as (27) below and (28) below (these represent examples of mandative subjunctive). The lack of inflection here is not surprising: Fulk (2012: 72), for example, refers to the early northern loss of <e> in the singular and <en> in the plural on verbs in the subjunctive, with King (1997: 179) stating that in Scots, such verbs are uninflected, too (except for the use of the verbs *be* and *have*); however, she also adds about the subjunctive in the north,

“[s]ubjunctive mood could ... be said to be indirectly related to Scots and in the north in contexts where an indicative inflection would normally occur owing to the operation of the NPTR, but as an expression of subjunctive mood a \emptyset -morpheme is found instead”.

- (27) *Thurgh vertu of þe haly gaste,
Through virtue of the holy ghost,
So þat he gude werkes bigin,
So that he good works begin,
And fulfils penance for his sin
And fulfils penance for his sin*
(1350-1420 *The Northern Homily Cycle* P III p. 137; *HCET*)

‘through [the] virtue of the Holy Ghost, so that he [the man] begins good works, and fulfils penance for his sin’

- (28) *And þe lord~ think~ spedfull þ=t= þ~ lord~
And the lords thinks expedient that the lords
begyn to syt on þe monu~day þe viij
begin to sit on the Monday the viij
day of noue~ber in Edinburghe
day of November in Edinburgh*
(1420-1500 *The Acts of the Parliaments of Scotland 1124-1707* p. 46.C2; *HCOS*)

‘And the lord[s] think it appropriate that the lord[s] begin to be in session on the Monday, the 8th day of November, in Edinburgh’

When used in the imperative, *beginnen* is only found in the NME data set, as in (29) below. In this example, this verb is unmarked, because it illustrates the use of this function in the singular (covert second person subject), where there was no inflectional ending in the north (Fulk 2012: 72) (and in Scots, as proposed by King 1997: 179).

- (29) *After: bigyn þi matyns; bot first: crosse þi
After: begin thy matins; but first: cross thy
lippis & sai: Domine, labia m[ea] a[peries]
lips & say: Domine, labia m[ea] a[peries]*
(1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 146; *ICoMEP*)

‘After: begin your office of matins [religious service]; but first: cross your lips and say: *Domine, labia m[ea] a[peries]* [*O Lord, open my lips*]’

However, in the plural, the inflectional ending of verbs in the imperative was <es> in the north (Fulk 2012: 72) (or <(i)s> in Scots, as in *kepys [y]ow fra disparying*, as argued by Macafee (1992/93: §8.3) and King (1997: 179), unless the subject was an adjacent pronominal pronoun, as in (30) below (overt first person subject), in which case the inflectional ending was either <e> or <∅>, as demonstrated by Cole (2014: 159), “in the northern Middle English present-indicative and imperative paradigms adjacent plural personal pronouns triggered verb forms ending in *–e* (later *–∅* by regular sound change), while all other subject types occurred with verb forms in *–s*.” In this respect, the imperative behaved like the present indicative.

- (30) *Begynne* *we* *at* *Cownsel,* *for* *þarof* *es*
 Begin we at Counsel, for thereof is
maste nede at þe begynnyng of owre werkes
 most need at the beginning of our works
 (1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 45;
ICoMEP)

‘We begin at counsel [one of the gifts of the Holy Spirit], because there is a great need for it at the beginning of our works’

3.3.2.3 Present Participle, Past Participle and Infinitive

Unlike the verb *gan* and its variant *can*, which occur in the finite form only in the NME and ESc data sets, the verb *beginnen* is additionally found in non-finite functions in these data sets, namely the present participle, the past participle and the infinitive. In such cases, the presence of endings also appears to be in line with what is expected of lexical verbs at the time, as discussed here. According to Table 33 and Table 34, forms of *beginnen* in the present participle end in <and(e)> in both data sets in both periods, as in (31) below, (32) below and (33) below. However, there is also a solitary instance of this verb in <yng> in the ESc data set in Period II, as in (34) below. In this example, *BEGYNNYNG* appears to have an appositive function, modifying the NP *ALL & HAILE Our LANDis WT Yar ProFITABIL PerTENENCE*. The predominant occurrence of *and* inflections on the verb *beginnen* in the NME and ESc data sets is, therefore, not surprising, as Macafee (1992/93: §8.5), King (1997: 180) and Fulk (2012: 73) state that *and* was the usual present participle ending on verbs in the north. Gardela (2014: 172-173), for example, found that

endings in *ing* began to occur on verbs in appositive and adjectival functions between 1350 and 1420 in his database of NME texts, but between 1420 and 1500 in his selection of ESc texts. Therefore, the occurrence of <yng> on this verb in the ESc data set in Period II might perhaps constitute one of the early instance of this variant replacing <and>, especially as Macafee (1992/93: §8.5) states that <ing> increasingly replaced <and> in OSc in the 16th century, suggesting that this process began earlier, as confirmed by Gardela (2014).

Table 33

*Forms of the Present Participle of **beginnen** in the NME Data Set in Period I and Period II*

		<i>beginnen</i>		
		Form	<i>N</i>	%
Period I	<bygynnande>		1	100
	<bigynand>		1	100

Notes: *N* – Observed absolute frequency
% – Per-cent of occurrence of *beginnen* per period

Table 34

*Forms of the Present Participle of **beginnen** in the ESc Data Set in Period I and Period II*

		<i>beginnen</i>		
		Form	<i>N</i>	%
Period I	<begynande>		1	20
	<begynnand>		2	40
	<begynnande>		2	40
Period II	<begynand>		4	17
	<begynande>		1	4
	<begynnand>		14	61
	<begynnande>		3	13
	<begynnyng>		1	4

Notes: *N* – Observed absolute frequency
% – Per-cent of occurrence of *beginnen* per period

- (31) *ffor [his] ioy is þe taste of þe blise of
 for [this] joy is the taste of the bliss of
 hewen, þe wilk is endles meyde of
 heaven, the which is endless reward of
 a deuote saule, þygynnande here.
 a devote soul, beginning here.
 (1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 167;
ICoMEP)*

‘for [this] joy is the taste of the bliss of Heaven, which is an eternal reward for a devoted soul, coming into existence here.’

- (32) *TO TerME DUYRANDE YE TYME OF YE
 to term during the time of the
 FORSAYDE TAKE YE TerME OF YE SAYDE
 foresaid tack the term of the said
 ERILIS ENTRE IN YE SAYDE TAKE
 earls entry in the said tack
BEGYNNANDE AT YE FEST OF SAYNT
 beginning at the feast of Saint
 MARTYNE
 Martin
 (1392 *Lennox Charters and Letters* 168; *LAOS*)*

‘to term during the time of the foresaid tack [i.e. a lease granting a tenancy of land or property], the term of the said earl’s entry in the said tack beginning at the feast of Saint Martin’

- (33) *jN YE FIRST BEGYnNAND AT YE
 in the first beginning at the
 DOUBFURDE OF LUCHLAND AND SYN
 doubford of Luchland and syne
 STREKAND VP TIL A BLINDFURDE BENETH
 streeking up to a blindford beneath
 YE SILUerFURDE
 the silverford
 (1446 *Registrum Episcopatus Aberdonensis* 174; *LAOS*)*

‘in the first, beginning at the doubford of Luchland and then extending to a blindford beneath the silverford’

- (34) *ALL & HAILE Our LANDis WT Yar
 all & whole our lands with their
 ProFITABIL PerTENENCE BEGYNNYNG AT YE
 profitable pertinence beginning at the
 BRIGENDE OFF KYLWYnNYnG ON YE WEST
 Bridgend of Kilwinning on the west*

SYDE OF YE WATer
 side of the water
 (1487 *Fraser Charters* 415; *LAOS*)

‘all of our lands with their valuable addition beginning at the
 Bridgend of Kilwinning on the west side of the water’

The past participle forms of *beginnen* are provided in Table 35 and Table 36 for the NME and ESc data sets, as in (35) below and (36) below, respectively.⁶⁰

Table 35

*Forms of the Past Participle of **beginnen** in the NME Data Set in Period I and Period II*

		<i>beginnen</i>		
		Form	<i>N</i>	%
Period I		<begonn>	1	14
		<begon>	1	14
		<begune>	1	14
		<begunne>	1	14
		<bigunne>	2	29
		<bygun>	1	14
Period II		<begun>	1	17
		<begon>	5	83

Notes: *N* – Observed absolute frequency
 % – Per-cent of occurrence of *beginnen* per period

⁶⁰ Although the verb *have* is the dominant auxiliary for the perfect in OSc (Macafee 1992/93: §8.9), it appears that non-finite *beginnen* can still occur with *be* in the ESc data set, where it occurs as an inflected form <beis>, functioning as an alternative to the present indicative form <is> (third person singular) with a future reference in a subordinate clause with *gif* (cf. *DSL* s.v. *be* v. 2b).

Table 36

*Forms of the Past Participle of **beginnen** in the ESc Data Set in Period I and Period II*

		<i>beginnen</i>		
		Form	N	%
Period I				
		<begunyn>	1	25
Period II		<begwn>	1	25
		<begunnyn>	1	25
		<vnbegownyn>	1	25

Notes: N – Observed absolute frequency
% – Per-cent of occurrence of **beginnen** per period

(35) *for to strengþe þe in þi gode worching*
for to strengthen thee in thy good working
þat þou hast be-gunne,
that thou has begun,
 (1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 267;
ICoMEP)

‘in order to strengthen you in your good working that you have begun,’

(36) *THAT HE HAS MOVIT AND BEGUnNYN AND*
that he has moved and begun and
BE ALL HIS POWERE
by all his power
 (1482 *Edinburgh, National Archives of Scotland, W Register House,*
PA 2/2, f61r - f64v 9519; LAOS)

‘That he has moved and begun, and by all his power’

In terms of inflectional suffixes of **beginnen** in the past participle, Fulk (2012: 74) proposes that in the north, “-en is retained in the strong [past participle]”, although he does not state how long it is used there. However, this ending does not appear to apply to the past participle forms of **beginnen** in the NME data set, as they generally occur with no inflection there, except for the final <e>, as in (35) above, which might be seen as a reduced form of the past participle inflection of strong verbs. Unlike in the NME data set, the past participle form of **beginnen** is still found with the

inflectional ending <yn>, i.e. a variant of <en>, in the ESc data set, as in (36) above and (37) below, suggesting that the strong past participle ending is still retained on this verb towards the beginning of the MSc period (the prefix <vn> in (37) below appears to be derivational, rather than inflectional, conveying here the idea of negation, rather than completion).⁶¹ In terms of the stem, forms of *beginnen* in the past participle occur with the orthographic variants <u> and <o> in the NME and ESc data sets, as in (35) above, (36) above, (37) below and (38) below. However, it is likely that in the NME data set, the realisations of both <u> and <o> were either /ʊ/ or /u:/ before a nasal consonant, rather than /ɔ/ and /ɔ:/, as suggested by the spelling of the form of *beginnen* in the past participle with <o>: a) the former Class III of OE strong verbs was characterised by the occurrence of the ablaut vowel /u/ with the distinguishing nasal /n/ (Hogg 2002: 59-60, Hogg & Fulk 2011: §6.51); and b) ME scribes adopted a convention (probably as a result of the French influence) of substituting <u> with <o> when /ʊ/ occurred next to <m>/<n> and <u>; this is because <u> (formed of two minims) was likely to cause confusion when using in combination with <m>/<n> and <u> (Fulk 2012: 27). In relation to <VNBEGOWNYN>, as in (37) below, it is probable that <ow> was realised as /u:/ in this form in the ESc data set, especially that the French spelling <ou>, including <ow> at the time, for /u:/ was used in ME by the 14th century (Fulk 2012: 27). Nevertheless, the remaining instances of *beginnen* in the past participle in the ESc data set are realised with either <u> or <w>, representing /u/, rather than <o>, despite the presence of the neighbouring <n>.

(37) *GIF* *IT* *BEIS* *VNBEGOWNYN* *TO* *BIG* *ON* *IN*
if it beis unbegun to bigg on in
THE *MEYN-TYME* & *GIF* *THAI* *INTer* *TO*
the meantime & if they enter to
BIG *IT*
bigg it
(1457 *Aberdeen Council Register II*, 22 Nov 1457 319; *LAOS*)

‘if it be unstarted to build on [the land] in the meantime, and if they enter to build it’

⁶¹ Macafee (1992/93: §8.7) proposes that the prefix <i> conveying the idea of completion and found on verbs in the past participle in SME as a relic of the OE prefix <3e> survives in OSc verse as an Anglicism.

- (38) *when þe psalm was begon in a medull voyce,*
when the psalm was begun in a middle voice,
& all þe monkis felowd after
& all the monks followed after
 (1420-1500 *Alphabet of Tales* pp. 87-88; *ICoMEP*)

‘when the psalm was begun in a voice of middle range, and all the monks followed after’

Forms of the verb *beginnen* are also found functioning as an infinitive in the NME and ESc data sets, as presented in Table 37 and Table 38, respectively (no distinction has been made here between different types of infinitive for ease of reference – different types of infinitive found with the verb *gan* and its variant *can*, as well as the verb *beginnen* in the NME and ESc data sets will be discussed in Chapter Four). In this function, forms of *beginnen* generally have no inflectional ending, as in (39) below and (40) below. However, some instances of forms of this verb in both data sets in Period II also show evidence of the final <e>, as in (41) below and (42) below, which could be the remnant of <en> used for the infinitive. This constitutes 11% and 50% of all forms of this verb in this function in the NME and ESc data sets, respectively, in this period.

Table 37

*Forms of the Infinitive of **beginnen** in the NME Data Set in Period I and Period II*

		<i>beginnen</i>		
		Form	N	%
Period I	<begyn>		1	33
	<bigin>		2	67
Period II	<begyn>		6	67
	<begyne>		1	11
	<begynn>		1	11
	<begynne>		1	11

Notes: N – Observed absolute frequency
 % – Per-cent of occurrence of *beginnen* per period

Table 38

*Forms of the Infinitive of **beginnen** in the ESc Data Set in Period I and Period II*

		<i>beginnen</i>		
		Form	N	%
Period I		<begyn>	1	100
		<begin>	2	20
Period II		<begyn>	3	30
		<begyne>	5	50

Notes: N – Observed absolute frequency
% – Per-cent of occurrence of **beginnen** per period

- (39) *Gude werkes gladly we suld bigin,*
 Good works gladly we should begin,
þat vnto þat welth might vs win,
 that onto that bliss might us get,
 (1350-1420 *The Northern Homily Cycle* P II p. 206; *HCET*)

‘We should gladly begin good works that might get us happiness,’

- (40) *PerTENAND TO YE TEmPIL OF SAYNT*
 pertaining to the temple of Saint
jON TO BEGYN AT YE EST HALF AT YE
 John to begin at the east half at the
BURN OF YE ARDACH AS IT DESCENDIS IN
 burn of the Ardach as it descends in
WATIR OF DEE
 water of Dee
 (1400 *Registrum Episcopatus Aberdonensis* 178; *LAOS*)

‘pertaining to the temple of Saint John to begin at the east half at the Ardach burn as it descends into the water of Dee’

- (41) *BE DECESs OF SCHIR WILLIAM*
 by decease of Sir William
HESLIHOIP LAST POSSESSOUR YAN WT FOUR
 Heslihoip last possessor then with four
MONETH TIL RYNE AND TIL BEGYNE FRA YE
 month to run and to begin from the
SAID TERME OFF WIT-SONDA [sic]
 said term of Whitsunday
 (1433 *Registrum Episcopatus Brechinensis* 1410; *LAOS*)

‘upon [the] death of Sir William Heslihoip, the last occupier with four

months to have authority and to begin from the said term of Whitsunday’

(42) *AND THE SAID V PerSONES SAL BEGYNE TO*
and the said v persons shall begin to
SITT APON THE SAID MATerEz TO-MORN
sit upon the said matters to-morn
(1459 *Aberdeen Council Register II*, 6 Nov 1459 p. 378; *LAOS*)

‘and the said five persons will start deliberating upon the said matters tomorrow’

The findings of this study in relation to the presence of this ending on the verb *beginnen* in infinitive also echo those on the subject provided in some sources, but only in terms of the type of inflection, rather than the timing. For example, Fulk (2012: 75) lists \emptyset and *e* as endings on strong verbs in the infinitive in the north, while King (1997: 179) states that “[a]part from some relics of final schwa in very early Scots, the infinitive was unmarked in Older Scots.” Therefore, while it is possible that forms of this verb in infinitive discussed here are used with <e> as a remnant of the infinitival ending <en>, the final <e> may also be indicative of orthographic variation.

3.4 Morphological Processes Involved in the Grammaticalization of *Gan*, *Can* and *Beginnen* in Northern Middle English and Early Scots

The occurrence of *gan* and *can* in finite form only in the NME and ESc data sets, combined with that of *beginnen* in both finite and non-finite forms in these data sets, suggests that these verbs differ with respect to their morphological paradigms, in view of what we know about grammaticalization and the development of invariant forms: a) *gan* and *can*, as well as *beginnen* participate in the global language change characterised by the attrition of inflectional endings and the singular and plural dichotomy in the root vowel in the preterite, affecting the category of verb in the transition from OE to ME; b) *gan* and *can* are also characterised by decategorialization and phonological reduction, suggesting a process of auxiliatio; and c) further developments involve the emergence of *begouth* as a variant of *beginnen* as a result of the process of analogical change. Below I discuss the morphology of the above-mentioned verbs in ME and Scots, including the NME and

ESc data sets, respectively. I then provide information on what decategorialization is and why it is relevant to this study. I subsequently discuss analogy in relation to the variant forms *couth* and *begouth*. I end this section with a discussion of whether or not *gan* and *can* undergo phonological reduction.

3.4.1 Morphological Behaviour of *Gan* and *Can* in Middle English Indicative of Grammaticalization

Since *gan* and *can* occur in the preterite in ME and Scots, suggestions have been made in the literature that this verb and its variant are morphologically fixed in the third person singular (i.e. finite form when in auxiliary use). Such statements should, therefore, be understood as indicative that *gan* and *can* are subject to the process of auxiliation (I discuss the link between morphological fixation and grammaticalization below). For example, Denison (1993: 322) refers to “the virtually fixed form *gan*, sometimes *can*”, whereas Ogura (1997: 403) argues that this verb “was found petrified in Middle English to the preterit singular form *gan* (or *gon*) and was used as an auxiliary that represented the preterit tense of the infinitive following”. However, in addition to statements in the literature highlighting the fact that both *gan* and *can* are morphologically fixed in the third person singular, a small number of sources refer to the presence of plural forms of this verb and its variant. Brinton (1996: 75, 297-310), for example, lists both singular and plural forms of *gan* from Chaucer’s *Troilus and Criseyde*, but she does not explicitly refer to the latter ones in the discussion of functions of *gan* in the above text. In a similar fashion, the *MED* (s.vv. *can* v.; *ginnen* v.), the *OED* (s.vv. *can* v.²; *gin* v.¹) and Denison (1993: 322) provide plural forms of this verb in ME, even in auxiliary use, as in (43) below. However, such forms may be restricted to southern varieties of ME, as is the case for this example, which is found in MS Harley 2253, the dialect of which is described as “HRF [Herefordshire]” (*eLALME* LP9260). The attrition of marking of inflectional endings to differentiate between the singular and plural forms in the preterite happened differently in various dialects of ME, with northern dialects being prone to the loss of inflection earlier than southern ones, as discussed in Section 3.2. Furthermore, the occurrence of *gan* and *can* in singular and plural forms in the preterite alone does not, however, suggest that this verb and its variant do not

undergo the process of auxiliation. Even the OE predecessors of modal (auxiliary) verbs in ME, especially at the beginning of the period, were inflected, i.e. *þu cannst*, *we cunnan*, *we cuðon*, except in the third person singular in the present tense (Lightfoot 2006: 94, cf. Warner 1993: 101).

- (43) *Payenes þer connen* [Trinity College, Cambridge MS R.3.8
gunne]
 Pains there did
aryue
 arrive
 (c1325 *King Horn* 187; *MED* s.v. *can* v.)

‘Pains arrived there’

According to the traditional view, morphological fixation is one of the diagnostics of the process of grammaticalization taking or having taken place (Hopper 1991: 17-36). For example, when discussing the development and function of *gan* in ME, Brinton (1996: 81) observes that one of the properties characterising this verb is “morphological ‘fixation’ ... since *gan* becomes fixed in the third-person preterite form, most often singular (even being confused in its northern form, *can*, with the modal verb)”, but in addition to the other aspects of the process of grammaticalization having an effect on the auxiliary status of this verb in Chaucer’s *Troilus and Criseyde*: a) syntactic fixation, which results in the verb *gan* being used with the plain infinitive, rather than the *to*-infinitive and being more adjacent to its infinitive in later texts; b) an initial loss of scope, whereby this verb came to modify a verb, instead of taking a *to*-infinitive complement; c) a subsequent increase in scope, in which case *gan* has a pragmatic meaning, i.e. textual, this verb subsequently functioning as a discourse marker in a range of contexts⁶²; d) divergence, whereby the verb *gan* functions as a marker of ingressive aspect in some contexts, while assuming textual, and, subsequently, interpersonal and metrical, functions in others (auxiliary use). These functions are potential diagnostics of grammaticalization. For example, the ingressive one is concerned with Hopper’s (1991) one of his five principles, namely persistence, whereas the textual one, but

⁶² However, an increase in scope is sometimes said to be counter to grammaticalization. See Fischer (2007: 259-322) for a discussion on this issue.

also interpersonal and metrical ones, are more to do with semantic changes that Brinton (1996) identifies in grammaticalizing forms; and e) renewal, whereby the verb *beginnen* assumes the ingressive function initially associated with the verb *gan* before its auxiliation, resulting in the loss of *gan* in this role. With these characteristics, Brinton (1996), therefore, shows that *gan* undergoes a change from lexical to grammatical.

The above proposal about the morphological fixation appears plausible, especially as it is considered within the context of the theory of grammaticalization. In particular, Brinton's (1996: 297-310) claims are strongly supported by evidence: 93% of all instances of the verb *gan* listed in her Appendix E are fixed in the third person singular preterite form <gan>, with only the remaining 7% being either <gonne> or <gonnen> (Brinton 1996: 297-301). In this respect, the occurrence of the verb *gan* almost exclusively in the third person singular preterite suggests that not all inflectional categories usually used to group word forms into paradigms, such as the present tense, the subjunctive and the present participle amongst others, apply to this verb, in accordance with the requirement of syntactic rules. However, while the evidence of *gan* occurring in the preterite form only, together with the other characteristics typical of grammaticalization, points to this verb undergoing the process of auxiliation, Brinton (1996: 81) also suggests that the loss of inflectional variability on this verb with respect to the number and person in the preterite could be as a result of the process of grammaticalization. Nevertheless, as discussed in Section 3.2, the paradigm of the verb was undergoing a number of changes concerned with the loss of inflectional morphology in general in the transition from OE to ME. Furthermore, these changes have in turn resulted in new category distinctions, whereby the surviving verbs of the preterite-present class came to be distinguished as a new grammatical category (Lightfoot 2006: 95). The process responsible for a shift in categorical relations is known as decategorialization, as discussed below.

3.4.2 Decategorialization

One can speak of decategorialization when a form loses its properties (regardless of whether morphological or syntactic) that would otherwise identify it as a full member of a lexical category. For example, in relation to morphological criteria, which can constitute a criterion in class membership, verbs in the earlier varieties of English are typically classified through a number of properties, for example tense, aspect, modality, person and number agreement. Hopper & Traugott ([1993] 2003: 106) argue that decategorialization involves changes “from a morphologically ‘heavier’ unit to one that is lighter, that is, from one that tends to be phonologically longer and distinct (e.g. with stressed vowels) to one that tends to be less distinct and shorter.” It is characterised by a tendency for prototypical members of a specific category to become less prototypical in terms of their distribution, but also in, at least, one of their uses. This kind of approach to decategorialization, i.e. involving “the loss of morphosyntactic properties, such as inflections and/or the ability to take articles and modifiers” (Norde 2009: 72), is, therefore, considered to be from a formal perspective. An example of this kind of decategorialization involves lexical verbs undergoing a process of grammaticalization to become modal (auxiliary) verbs in ModE, where instances **he cans* and **he wills* are seen as ungrammatical. However, decategorialization as a shift from one category to another can be viewed from a functional perspective. Norde (2009: 72), for example, argues that in such cases, it entails, “loss of discourse autonomy” (cf. Hopper 1991: 30), where verbs, for example, no longer have an ability to provide information about new events, as in (44) below.

- (44) *They saw the Northern Lights* (referential)
 Seeing that you have declared bankruptcy, you can hardly make any
 new investments (grammaticalized, non-referential) (Norde 2009: 72)

The process of decategorialization from a formal perspective, Hopper & Traugott ([1993] 2003: 106-107) argue, is not about the decay or deterioration of that form, but a shift from one kind of role to another. In other words, it is possible to speak about a cline of grammaticalization, whereby a starting point for a lexical item is a

full category, with intermediate points being described in terms of a loss of morphological structures associated with this full category. Hopper & Traugott ([1993] 2003: 106f), therefore, define decategorialization as “the tendency for relative prototypical members of Noun, Verb, and Adjective categories to become less prototypical in their distribution”. This tendency is characterised by a loss of morphological and syntactic properties, as a result of “a functional shift from one kind of role to another in the organization of discourse”. In this respect, this emphasises the idea of gradualness, whereby the kinds of changes undergone by the grammaticalizing element are a step-by-step process (references here). This process, ultimately, results in category shift, a view also adopted by a number of others (Norde 2009: 72). For example, a verb-to-affix cline described by Hopper & Traugott ([1993] 2003: 110) involves a number of points characterised by a change from a full verb > auxiliary > verbal clitic > verbal affix, with examples from a lexical verb to an auxiliary verb, including ModE ‘quasi-auxiliary’ *have to read a book*, to ModE ‘auxiliary’ *have had a book*. In this shift, as the process of grammaticalization advances, “verbs may lose such verb-like attributes as the ability to show variation in tense, aspect, modality, and person-number marking” (Hopper & Traugott [1993] 2003: 108).

However, decategorialization from the formal perspective may not necessary consist of a category shift (Norde 2009: 73). Lehmann (2004: 167), for example, uses the term “recategorialization”, viewing it as a type of change, but not in the sense of category change, whereby an element belonging to one category becomes recategorialised as, and associated with, another category in its morphological and syntactic context. For Lehmann (2004: 165-167), therefore, recategorialization is defined in two ways: a) the context of the grammaticalizing element is reinterpreted; and b) this type of change can involve changes within one category in a move from lexical to grammatical (cf. Breban 2010: 88-90). Diewald (1997) and Norde (2009) give an example of “the well-known semantic-pragmatic shift from deontic to epistemic modality, illustrated by German auxiliaries” (Norde 2009: 73), as in (45) below. Both argue that when occurring with epistemic meaning, some forms of German modal (auxiliary) verbs can be prevented from being formed, more

specifically the preterite forms of *wollen* in this example. Norde (2009: 73-74) then concludes that, “*wollen* [has] clearly become restricted paradigmatically when [it] came to be used in epistemic constructions, yet there is no shift from one category to another.” In other words, even though it is possible to observe a shift from deontic to epistemic in the meaning of this modal (auxiliary) verb in German, there is no decategorialization involved in this example in the sense of category change. There is a lack of decategorialization because we are dealing with an already grammaticalized element, although it is possible for an already functional element to become even more functional.

- (45) A. *Sie will im Büro gewesen sein*
 She will-Present in.the-Dat office been
 have
 ‘She has probably been in the office’
- B. *?Sie wollte im Büro gewesen sein*
 She will-Preterite in.the-Dat office been
 have (Norde 2009: 73; translation in the original)

Despite the above-mentioned problem with decategorialization not necessary involving a category shift, I retain here the term ‘decategorialization’, but restrict its scope to refer to the loss of morphosyntactic properties.

3.4.2.1 *Changes in the Category of Gan and Can in Northern Middle English and Early Scots*

The findings of *gan* and *can* in the NME and ESc data sets, as outlined in Section 3.3.1, show that this verb and its variant only occur in finite form (the preterite) in these data sets, as in (46) below and (47) below. I would, therefore, argue that the lack of non-finite forms, typical of other verbs at the time, including the verb *beginnen*, as outlined in Section 3.3.2, suggests that *gan* and *can* are subject to decategorialization, i.e. a shift to a functional category, resulting in morphologically heavier items becoming lighter (Hopper & Traugott [1993] 2003: 106) by losing their morphosyntactic properties (Norde 2009: 72), such as an ability to inflect. Further evidence for the decategorialization of *gan* and *can* in the NME and ESc

data sets is found when comparing the morphology of this verb and its variant to that of *ginnen* found in ME only. It will be recalled from Section 1.5.1 that *ginnen* occurs in finite and non-finite forms. For example, in (48) below, <gynneth> is found in the third person present tense, whereas in (49) below, <gynne> occurs in the infinitive, both instances of this verb being in absolute uses. Given that *ginnen* has a full morphology (finite and non-finite forms) and predominantly occurs with the *to*-infinitive, both properties being atypical of verbs otherwise characterised by auxiliatio (cf. Warner 1993: 103), and that it is also predominantly found with ingressive meaning, uses of this verb would (presumably) be deemed as non-auxiliary (cf. *MED* s.v. *ginnen* v. 3b, which argues that *ginnen* in the present tense found with the plain infinitive also has “weak auxiliary” uses). In contrast, with *gan* and *can* being in the preterite only, with the plain infinitive and non-ingressive meaning, this verb and its variant cannot be simply considered as the third person preterite forms of *ginnen*. Instead, what is initially one category of the verb is ultimately reanalysed as two: *gan* and *can* as auxiliaries, defined by the existence of the finite form (in addition to occurring with the plain infinitive and having a non-ingressive meaning), whereas *ginnen* as a lexical verb, distinguished by a range of finite and non-finite forms.

- (46) *When scho was flemyd that was so gent,*
When she was banished that was so gentle,
Ilke a lorde to hys lande es went,
Each a lord to his land is gone,
For sorow thaire hertes gan blede.
For sorrow their hearts began to bleed.
 (1420-1500 *Octavian* l. 289-291; *TEAMS*)

‘When she, who was so gentle, was banished, each lord went to their country, and their hearts began to bleed with sorrow.’

- (47) *Wit þat stan he laid in sling,*
With that stone he laid in sling,
Sua stalworthli he lete it suing
So strongly he let it swing
Pat in his frunt þat stan he fest,
That in his forehead that stone he cast,
Pat bath his eien vte can brest
That both his eyes out did burst
 (1350-1420 *Cursor Mundi* p. 438; *HCET*)

‘He put that stone in the sling, and let it swing so strongly that he cast it on the forehead, that both of his eyes came out’

- (48) *Heere gynneth the prologe in the boc of*
Here begins the prologue in the book of
Ecclesiastes.
Ecclesiastes.
(a1382 *Wycliffite Bible* (early version) *Ecclesiastes Prologue*; *OED*
s.v. *gin* v.¹ 2a)

‘Here begins the prologue of the Book of Ecclesiastes’

- (49) *Pe flode bigan to gynne, & klosed it* [the island]
The flood did begin, & closed it
aboute
about
(c1330 Robert Mannyng of Brunne *Chronicle* 77; *OED* s.v. *gin* v.¹ 2a)

‘The flood started and encircled the island’

Both *gan* and *can*, therefore, show no morphological properties in the NME and ESc data sets that would otherwise categorise them as members of a major lexical category of the verb (Lehmann 1985: 307; 1995: 132, Hopper & Traugott [1993] 2003: 107). Instead, they show morphological characteristics that are typical of auxiliaryhood. For example, Lightfoot (1979; 1991: 142 ff) makes the restriction of some verbs to finite forms only as a factor preconditioning the development of the class of modal (auxiliary) verbs. In a similar way, Warner (1993: 144-145) argues that the absence of some non-finite forms on modal (auxiliary) verbs is a striking feature, even in IME, making this a formal distinction from a wider class of verbs during this time. However, although he points out, “a rather fuller morphology [i.e. non-finite forms] for several of these verbs is recorded by late Middle English”, “this apparent development ... argues that [a modal group within an auxiliary group] none the less remained within the more general class of verbs” (Warner 1993: 144-145). In other words, although some modal (auxiliary) verbs have finite forms only, we are dealing here with the issue of category change, whereby some verbs, deemed to belong to a modal group within an auxiliary group, still have non-finite forms. Therefore, the evidence presented here shows that *gan* and *can* in the NME and ESc data sets are characterised by decategorialization; this verb and its variant no longer

belong to a major, lexically ‘open’ category in terms of their morphological properties, but a minor, relatively ‘closed’, category, which is similar to that of modal (auxiliary) verbs with respect to the lack of non-finite forms for, at least, some of them even in IME. This suggests rightward movement for *gan* and *can* in the NME and ESc data sets along the cline of grammaticality, as the process of grammaticalization advances.

3.4.2.2 *Loss of Inflectional Endings in the Preterite of Gan, Can and Beginnen in Northern Middle English and Early Scots*

Since *gan* and *can* occur in the preterite in the NME and ESc data sets, I only refer in this section to inflectional endings and root vowel alternations associated with this tense, comparing and contrasting this aspect of morphology of this verb and its variant with that of the verb *beginnen*. As discussed in Section 3.2, OE Class III strong verbs were marked in the preterite by having an inflectional ending, in addition to changes in the root vowel (ablaut) not only to distinguish between the preterite and non-preterite forms, but also to differentiate between the singular and the plural. However, during the transition from OE to ME, the category of the verb underwent changes in this tense: the attrition of inflectional endings to mark the singular and plural dichotomy and the levelling of the root vowel. This remodelling of inflectional morphology on the category of the verb in the preterite began in the north, including Scots, gradually finding its way into other dialects of ME (Wyld 1914: 205-206, Lass 1992: 130-132, Macafee 1992/93: §8.11, Fisiak 1996: 102, King 1997: 177, Fulk 2012: 74, 76). Given the above points, the lack of inflectional variability in marking of the singular and plural in the preterite (inflectional endings and root vowel alternations) on the verb *gan* and its variant *can*, as well as the verb *beginnen* in this tense in the NME and ESc data sets, as outlined in Section 3.3, can be accounted for by global language changes affecting all verbs (although there are some isolated instances of these verbs with the final <e>, these tend to be orthographic, rather than the remnants of the plural inflection). Further evidence in support of the attrition of inflectional endings and distinctions between the singular and the plural forms of the verb *gan* and its variant *can*, as well as the verb *beginnen* in the NME and ESc data sets as part of global language changes affecting all verbs

comes from the fact that these verbs show the lack of these features to the same degree in these data sets.

3.4.3 Analogy

Analogy is involved in change in general, especially, morphosyntactic change. Unlike reanalysis, which is responsible for underlying changes, affecting constituency, hierarchical structure, category labels and grammatical relations, analogy affects surface manifestations. In other words, while reanalysis is covert, analogy is overt. The concept of analogy goes back to Meillet ([1912] 1958: 130), who discusses this mechanism, stating that all regular forms of language can be described as analogous, as they are created on the basis of already existing constructions. This mechanism was seen as that of a proportion, whereby the alternation between *cat* and *cats* could be seen as extended to have *child* and **childs* (this form being found in child language). Despite its shortcomings, such a mechanism of proportions would also be useful in explaining analogies in the history of English, where, for example, *shoe* and *shoen* was perceived as realised as *shoe* and *shoes* on analogy with *stone* and *stones*, (Hopper & Traugott [1993] 2003: 64). But, this idea of analogy as a mechanism responsible for regularising irregularities in grammar does not account why one member of a group is selected as the model (Hopper & Traugott [1993] 2003: 64, cf. Kuryłowicz 1964: 38, Kiparsky 1992: 56). Furthermore, it views analogy as disjoined from what Meillet ([1912] 1958) calls grammaticalization and what is the result of what can now be called reanalysis. However, the neogrammarian view of analogy described above was first redefined by Kiparsky (1968), who viewed analogy as rule extension, i.e. generalisation or optimisation of a rule from a limited domain to a broader one (Hopper & Traugott [1993] 2003: 64). In this respect, analogy is best considered as generalization of a rule or construction, i.e. involved in paradigmatic organisation of changes in patterns of use.

The role of analogy in grammaticalization has long been recognised. But, as Traugott & Trousdale (2013: 37) demonstrate, this mechanism of change has reluctantly been recognised in some works on grammaticalization. For instance, they refer to

Haspelmath (1998) and Lehmann (2004), the latter of which distinguishes between “pure grammaticalization without analogy” and “grammaticalization with analogy”. However, the importance of analogy in the process of grammaticalization is emphasised by Anttila (2003) and Fischer (2007). The latter source also argues that the mechanism of analogy operates along both the pragmatic and syntagmatic axis; in this respect, according to Fischer (2007), the mechanism of analogy has a greater applicability than reanalysis. Traugott & Trousdale (2013: 37), on the other hand, recognise the importance of analogy in the process of grammaticalization. Referring to it as “analogization”, they see it as “a mechanism or process of change bringing about matches of meaning and form that did not exist before.” However, they also differentiate analogy or analogization from the process of analogical thinking, i.e. a motivation; in their view, “[a]nalogical thinking matches aspects of meaning and form; it enables, but may or may not result in change.” An instance of analogy in the history of English, where there is a formation of a new relationship between meaning and form, can be observed in the case of the *be going to* construction. Hopper & Traugott ([1993] 2003: 68-69) demonstrate that following the reanalysis of an auxiliary construction denoting the future, i.e. *be going to*, with a verb of activity, it is possible to observe the process of extension via the mechanism of analogy to all kinds of verbs, including stative verbs. Overall, therefore, analogy is concerned with paradigmatic change, i.e. surface collocations and/or patterns in use, whereas reanalysis operates along and involves linear and syntagmatic rule change or reorganisation (Hopper & Traugott [1993] 2003: 68).

From the items under investigation in this study, an example of analogical change can be observed in relation to the variant *can*, resulting in an emergence of *couth*, namely the preterite form of this variant, as in (50) below. Although no instances of *couth* are found in the ESc data set, both *can* and *couth* are used to refer to actions or events occurring in the past. The *OED* (s.v. *can* v.²) argues that “[c]hiefly” in Scots, *couth* frequently replaced *can* as a variant of *gan*, “in order to remove an apparent irregularity where what was perceived as a present stem of [*can* as a modal (auxiliary) verb] was used as a past tense of the periphrastic construction”, with Visser (1969: 1579-1580) also having instances of *couth* from NME and SME texts. Therefore, a new relationship is formed between meaning and form that did not exist

before, namely like *can*, the variant *couth* is also found in auxiliary use in ME and Scots.

- (50) *The croune, that Ihesu couth ber*
 The crown, that Jesus did bear
 (1375 John Barbour *The Bruce* III 460; *DSL* s.v. *can* v. 2b)

‘The crown, that Jesus bore’

Evidence of analogical change can also be seen on the verb *beginnen* in the ESc data set, where the variant *begouth* functions as an independent verb, i.e. without infinitival complementation and not in auxiliary use, as in (51) below and (52) below (in the latter, the text of the sample ends on *BEGANE*). There are two points to note about this variant: firstly, *begouth* is the outcome of analogical change in form, but not in meaning, with the preterite form *couth* as a variant of *can* (cf. Jamieson 1808 s.v. *begouth*, *begoude* pret. on the etymology of *begouth*); and secondly, both *begouth*, including other such variants, and *began* are irregular and are used in the preterite in the ESc data set, as in (51) below and (52) below. In this respect, the verb *beginnen* can be marked for tense (the preterite) with two different forms, a situation similar to that of *can* and *couth* in ME and Scots.

- (51) *AND SWA TIL IT ENTer IN GRIGis FURDE*
 and so until it enter in Grig’s ford
QUHAR WE BEGOUTH FIRST
 where we began first
 (1400 *Registrum Episcopatus Aberdonensis* 179; *LAOS*)

‘And so until it enter[s] Grig’s ford [in a river] where we began first’

- (52) *TO YE LANDis OF GARDYN AT FALLATY*
 to the lands of Gardyn at Fallaty
AT YE NORTH-EST DEVYDAND YE LANDis OF
 at the north-east dividing the lands of
GARDYN PRESCHOK AND DUnBERROW QUHAIR
 Gardyn Preschok and Dunberrow where
WE BEGANE
 we began
 (1434 *Arbroath Abbey* 322; *LAOS*)

‘to the lands of Gardyn at Fallaty in the north-east, dividing the lands of Gardyn, Preschok and Dunberrow, where we began’

In the case of *couth* and *begouth*, it is, therefore, possible to observe the mechanism of analogy; yet, although the force of analogy has a tendency to reduce the number of irregular forms over time, it also operates in the opposite direction here, with the irregular *couth* and *begouth* arising by analogy from existing irregular forms, i.e. the preterite form *can* as a variant of *gan* deriving from OE *onginnan* and OE and eME *aginnan*. Furthermore, the occurrence of the variant *couth* perhaps suggests a greater degree of grammaticalization of *can* as a variant of *gan*, in a similar way to the modal (auxiliary) verb *can* in ME and Scots.

3.4.4 Phonological Reduction

In addition to decategorialization, a grammaticalizing element also typically undergoes phonological reduction. Meillet (1912: 139; translation mine) encapsulates the whole process by referring to “l’affaiblissement de la forme” [the weakening of the form] and “affaiblissement de la prononciation” [weaker pronunciation], showing that not only is this process responsible for the loss of phonological substance, but it is also linked to morphology. Therefore, phonological reduction can be described as the loss of either individual segments or marked phonological features. For instance, Roberts & Roussou (2003: 224) provide an example of the Latin verb *habere*, which in Romance languages became reduced to a clitic and then an affix when functioning as a future/conditional marker; on the other hand, *that* in ModE is phonologically reduced when it functions as a complementizer, but not when its role in a clause is that of a demonstrative. The process of grammaticalization leading to a complete morphological independence of a grammatical item, but also that of phonological substance, can be illustrated using the cline of grammaticality in Hopper & Traugott ([1993] 2003: 7): content item > grammatical word > clitic > inflectional affix > phoneme > \emptyset (cf. Givón 1979: 213, Heine & Reh 1984: 113, Lehmann 1995: 25 on the notion of “grammaticalization channels”; Heine, Claudi & Hünnemeyer 1991b: 222 on the notion of “grammaticalization chains”). The above examples of *habere* and *that* also represent different degrees of grammaticalization, with the former being more grammaticalized than the latter. Furthermore, while Roberts & Roussou

(2003: 224), for example, maintain that “[t]hirteen out of eighteen of our cases of grammaticalization thus involve phonological reduction of some type”, the reminder having the status of “semi-functional”, a question needs to be raised whether phonological reduction is invariably part of all examples of grammaticalization.

As pointed out by Heine (1993: 106), Bybee, Perkins & Pagliuca (1994: 4-9), Lehmann (1995: 113), Roberts & Roussou (2003: 224-225) and L. Campbell (2013: 297) amongst others, phonological reduction is not confined to grammaticalization only, although it is central to this process or indicative that grammaticalization has occurred. For example, L. Campbell (2013: 297) argues, “[p]honological reduction processes apply to items of the appropriate phonological character generally in a language, not just to certain items which happen to be involved in processes of grammaticalization.” Crucially, it would be necessary to be able to separate instances of global language change affecting all verbs from those of grammaticalization, affecting only a select group of verbs. Roberts & Roussou (2003: 224-225), for example, maintain that although phonological change happens all the time, the kind of phonological change associated with the process of grammaticalization can be described as more “radical than standard phonological change”, exclusively affecting “grammaticalized elements”. In other words, phonological reduction as part of grammaticalization affects the fundamental nature of something, in this respect, being, in a sense, more localised. For example, during the grammaticalization of the *be going to*, *have got to*, *want to* and *be supposed to* constructions in ModE, it is possible to observe phonological reduction and coalescence of separate morphemes into one, namely *gonna*, *gotta*, *wanna* and *be suppose(d) to*, because the phrasal boundary between each separate morpheme is no longer there; therefore, the processes of phonological reduction and coalescence result in constructions that are no longer analysable into their grammatically individual parts (Leech, Hundt, Mair & Smith 2009: 105-107, cf. Hopper & Traugott [1993] 2003: 3). Finally, phonological reduction as part of grammaticalization may follow reanalysis (Hopper & Traugott [1993] 2003: 3), or it may either lead up to it or follow it (Lightfoot 1991: 171, Heine & Reh 1994: 17). Given this, Dér (2013: 168-169; emphasis in the original) argues,

[c]onsequently, phonological reduction that occurs in the course of a lengthy process of changes may or may not be related to the grammaticalization that

goes on simultaneously; this can only be claimed with confidence if it can be shown that the meaning of an item became (more) grammatical in nature **while** it was undergoing phonological reduction.

In relation to grammaticalization, Kenstowicz (1994: 640-642) and Roberts & Roussou (2003: 29, 37, 225-228) show that grammatical and functional elements, such as modal (auxiliary) verbs, which developed unstressed and reduced forms in the 16th century (Plank 1984), as well as complementizers and determiners, can phonologically contract. As a result, they are typically unstressed and ‘light’, or become a ‘prosodic dwarf’, to use Kenstowicz’s (1994) term. In terms of prosodic hierarchy, Roberts & Roussou (2003: 225, 224-229) observe that unaccented forms in ModE are always “subminimal”, whereas accented ones are bimoraic⁶³, “[e]very foot must be binary, that is, disyllabic or bimoraic, and so monosyllabic or monomoraic items cannot be feet, and therefore cannot be phonological words.” The term ‘subminimal’ is, therefore, used by Roberts & Roussou (2003: 225) to refer to phonologically-realised functional elements in relation to the prosodic system of the language, with a tendency to cliticise, the crucial information, therefore, being that subminimal forms cannot occur independently, i.e. as free forms.⁶⁴ After Roberts & Roussou (2003: 225) provide a list of “‘weak forms’” (accented and unaccented) of functional items, such as the auxiliary verbs *do* and *have* and the modal (auxiliary) verbs *can* and *will* in ModE amongst others from Gimson (1980: 261-263), they observe the following,

the unaccented form is subminimal (it is almost always monomoraic, i.e. containing something smaller than CVC or CV), the accented form is at least bimoraic, and nearly all the elements capable of being unaccented are functional. The unaccented forms are, of course, the usual ones in connected speech, unless the element in question is contrastively stressed. All the CVC unaccented forms contain schwa, which cannot be stressed in most varieties of English. (Roberts & Roussou 2003: 227)

Roberts & Roussou (2003: 227), therefore, note a correlation between functional and

⁶³ Kenstowicz (1994: 640) argues that monomoraic words are not present amongst lexical items in ModE, leading Roberts & Roussou (2003: 225) to define a minimal (lexical) word in ModE as bimoraic.

⁶⁴ However, caution is advised here, as Jucker (2002: 212) comments, “[i]n historical data ... there is no way of checking the actual pronunciation beyond the orthographic representation, which may or may not reflect the phonological reduction” (cf. Erman 1992: 219, Brinton 1996: 33, Jucker & Ziv 1998: 3).

lexical items: the former can have a restricted set of vowels and so can be unaccented, in addition to not being able to occur independently, whereas the latter are the opposite. For example, *do* functions as a lexical verb in *I do university administration every morning*, whereas in *Do universities serve any purpose?*, it functions as an auxiliary verb that can be reduced to either /də/ or /d/. In terms of this correlation between functional and lexical elements, Roberts & Roussou (2003: 228) justify why the process of grammaticalization may consist of phonological reduction, “[w]here a lexical element is reanalyzed as functional, the reanalysis must involve phonological reduction if functional categories are required to be subminimal”. They, therefore, imply here that the reanalysis of a lexical element leads to a functional one, but even if functional entails reduced, it does not necessarily mean that reduced entails or implies functional. In this respect, grammaticalizing items have a tendency to shorten. A similar observation is made by Hopper & Traugott ([1993] 2003: 154), who identify two tendencies in phonological attrition and selection, which accompanies the process of morphologization: a) a quantitative (“syntagmatic”) reduction: forms become shorter as the phonemes that comprise them erode; and b) a qualitative (“paradigmatic”) reduction: the remaining phonological segments in the form are drawn from a progressively shrinking set. However, while the former tendency can be applied to items undergoing phonological change in general, the latter one appears to be related to items undergoing grammaticalization. Therefore, in relation to the latter tendency, Hopper & Traugott ([1993] 2003: 155) argue that “[t]his smaller set of phonemes resulting [from this tendency] reflects the universal set of unmarked segments.” It includes such “apical” consonants as [n], [t] and [s], glottal consonants, such as [ʔ] and [h], as well as common vowels, such as [a], [u], [i] and [ə]. Below I show how certain properties of phonological reduction are applicable to the story of *gan* and *can* in ME and Scots, including the NME and ESc data sets.

Although auxiliation has been argued for the verb *gan* and its variant *can* in ME and Scots, the verb *gan* is only referred to in the literature as an “aphetic” form derived from OE *onginnan* and OE and eME *aginnan* (*MED* s.vv. *aginnen* v.; *ginnen* v., *OED* s.vv. *gin* v.¹; *ongin* v., Visser 1969: 1372). For example, Mossé (1952: 103)

argues that ME lexical base *ginnen* is “simply the equivalent, by the loss of the syllable, of *agin(ne)* ‘begin’ (< OE *onginnan*)”, whereas Brinton (1981: 156; 1996: 67) refers to ME *gan* as “the aphetic form” of OE *onginnan* (> *aginnan*). Although it is generally accepted in the literature that *gan* derives from OE *onginnan* and OE and eME *aginnan*, as outlined above, some sources suggest that this verb derives from the etymologically related OE *beginnan*. For example, Macafee (1992/93: §8.13) refers to *gan* in Scots as simply “a reduced form of *began*”, perhaps because two compounds, i.e. OE *be-* or *bi-ginnan* and OE *on-ginnan*, derive from the original Germanic verb **ginnan*. Similarly, the *OED* (s.vv. *begin* v.¹; *gin* v.¹) argues that *gan* is an “[a]phetic form of *begin*”, adding that “in early instances perhaps rather of *ongin*”. Although most sources agree that *gan* derives from *onginnan*, the statements about the origins of this verb in the *OED* suggest a lack of complete certainty about the etymology of this verb; it also implies some kind of split in terms of the grammatical status of both *gan* and *beginnen* in ME.⁶⁵ Overall, the above sources note that both *gan* and *can* occur as reduced or “aphetic” forms from the beginning of the ME period onwards. Although they do not specify that this reduction is the result of the process of grammaticalization, it is in combination with the other processes and mechanisms operating on the verb *gan* and its variant *can* in ME and Scots that this is perhaps the case. However, as highlighted further on, in the transition from OE to eME, it is also possible to observe the weakening of prefixes on compound verbs, such as OE *onginnan* and OE and eME *aginnan*.

In the NME and ESc data sets, both *gan* and *can* occur as aphetic forms, too. I show here that this is due to two separate, but related, processes of phonological reduction as part of grammaticalization. The first one results in a loss of the initial syllable affecting a wider range of verbs in OE and eME (given the limited availability of literary material written in Northern England before the beginning of the ME period and the lack of it in Pre-literary Scots to warrant a full lexico-grammatical analysis, I use the material available from OE and eME). This process affected a number of prefixes that were in operation at the time, such as *a*, *be*, *for*, *ge*, *of*, *to*, *on*, *purh* and

⁶⁵ Sims (2014: 75-76), on the other hand, argues that ModE *gan* is a “phonetically reduced” form of ModE *begin*; She concludes that *begin* “is susceptible to phonological reduction”, suggesting the re-appearance of phonologically reduced *gan* as a variant of *begin* in ModE.

ymb (Los, Blom, Booij, Elenbaas & van Kemenade 2012: 195). Van Kemenade & Los (2003: 104) observe that “prefixes in Old English ... are in an advanced state of grammaticalization”. Therefore, as an instance of grammaticalization, the *on* prefix, found frequently, but not exclusively, in stressed position on verbs underwent a weakening process (*OED* s.v. *an-*, *prefix*¹, de la Cruz 1975, Hiltunen 1983, Brinton 1988, as cited in Sims 2008: 119, cf. Sims 2014). As a result of this weakening, this prefix was subsequently realised as either *a* or *an* (the *a* prefix was used before consonants, whereas *an* before vowels), eventually becoming a reduced form used in unstressed position (*OED* s.v. *an-*, *prefix*¹). Quoting Hiltunen (1983: 92), Los, Blom, Booij, Elenbaas & van Kemenade (2012: 195-196) claim that both the *on* and *a/an* prefixes amongst others were still found in eME, but in much lower frequencies, almost completely disappearing at the beginning of the ME period.⁶⁶ As a result of this weakening process, it is possible to observe change from OE *onginnan*, as in (53) below, to OE and eME *aginnan*, as in (54) below, and, eventually, to ME *ginnen*, as in (55) below.

- (53) *Min child i blisse sone onngann To blissenn*
 Mine child in bliss soon began to rejoice
i min wambe
 in mine womb
 (?c1200 *Ormulum* (Jun 1) 2801; *MED* s.v. *aginnen* v.)

‘My child soon began to rejoice in my womb’

- (54) *He lokien agon in læðest alre monne*
 He look did in hateful all men
 (c1275 (?a1200) *Lazamon Brut*; *Caligula* MS A. IX 30831; *MED* s.v. *aginnen* v.)

‘He looked at the most hateful of all men’

- (55) *Heer gynmeth the prolog in the booc of*
 Here begins the prologue in the book of
Wisdam ... Here eendith the prolog ... Heer

⁶⁶ The phonological reduction of the prefixes is one of other processes typical of grammaticalization being under way. Hiltunen (1983: 94), for example, shows that the weakening of the prefixes was characterised by semantic bleaching, too. As a result of this, prefixes could no longer express their original meanings, for instance locative, aspectual and intensifying. This, in turn, appears to be in line with changes in relation to word order, namely the tendencies towards more analytical and more fixed SVO word order (Hiltunen 1983: 101).

Wisdom ... Here ends the prologue ... Here
gynneth the booc
begins the book
(a1382 *Wycliffite Bible Prologue to Wisdom* (Douce 369 MS) 85;
MED s.v. *ginnen* v. 1)

‘Here begins the prologue of the book of Wisdom ... Here ends the
prologue ... Here begins the book’

However, the weakening of both the *on* and *a/an* prefixes on compound verbs like OE *onginnan* and OE and eME *aginnan* appears to coincide to an extent with the emergence of the aphetic form *gan* in auxiliary use in ME and Scots. In this respect, phonological reduction that takes place in relation to the weakening of the above prefixes may or may not be related to the grammaticalization of the verb *gan* and its variant *can* in the NME and ESc data sets. For example, van Kemenade & Los (2003: 104) point out the loss of prefixes is advanced in OE, and Sims (2008; 2014) and Ogura (1997) elaborate on this in terms of particular dialects. Sims (2008: 105; 2014: 64) finds instances of *onginnan*, but no instances of the reduced form *aginnan*, in the Northumbrian and Mercian versions of *The Gospel of Saint Matthew* in OE, with Ogura (1997: 415) having only one instance of *aginnan* in the Northumbrian version of *The Gospel of Saint Luke*. Furthermore, the *on* and *a/an* prefixes are still found in IME, although, admittedly, these could constitute isolated cases, especially from SME. For example, the *OED* (s.vv. *agin* v.; *ongin* v.) has instances of *onginnan* and *aginnan* (in ingressive sense) from as late as a1300, as in (56) below, and c1400, as in (57) below, respectively. In contrast, early instances of *gan* are recorded by the *MED* and the *OED* to be from the beginning of the 13th century, as in (58) below, whereas those of *gan* and *can* in the NME and ESc data sets are dated 1350-1420 and 1375-1420, respectively, both referred to here as Period I, given the limited textual evidence in the Northumbrian dialect of OE and the lack of textual evidence in allowing a lexico-grammatical analysis in Pre-literary Scots (Craigie 1931: vii, Macafee 1992/93: §1.4, King 1997: 156, Scott 2004: 214, Cuesta & Silva 2008: 50-51).

(56) *Bi-þenc* *þe, mon, ant* *up* *aris* *of* *slovþe,*
Bethink you, man, and up arise of sloth,
an-gin to *worche* *god.*

be-gin to worship God.
(a1300 *English Lyrics 13th Century* 82; *OED* s.v. *ongin* v.)

‘Bethink you, man, and arise out of sloth, begin to worship God.’

(57) *Nov agynneb þe oþere partye*
Now begin the other party
Of Alisaunders dedes hardye.
Of Alexander’s deeds hardy.
(c1400 *Kyng Alisaunder* (Laud MS) l. 4747; *OED* s.v. *agin* v.)

‘The other party now begin[s] to encourage Alexander’s deeds.’

(58) *He gann þennkenn off himm self*
He did/began to think of him self
(?c1200 *Ormulum* (Burchfield Transcript) l. 3274; *OED* s.v. *gin* v.¹
1a)

‘He did/began to think about himself’

Although the weakening of prefixes occurs in OE on compound verbs such as OE *onginnan* and OE and eME *aginnan*, it is also possible to observe radical changes in terms of the grammatical status of these two verbs in OE already. Warner (1993: 117), Los (2000; 2005: 88-99) and Sims (2008; 2014) argue that *onginnan* (and *aginnan*) displays auxiliary behaviour (semantic bleaching) when it is used with the plain infinitive, in addition to having a function in a verse text (cf. Los 2000; 2005: 88-99 on the status of OE *beginnan* in Ælfric). Such changes in the grammatical status of OE *onginnan* and OE and eME *aginnan* would suggest that the monosyllabic form *gan* in the NME and ESc data sets would be the outcome of phonological reduction that has already happened, as part of the process of grammaticalization from *on* > *a* > \emptyset . This could only be maintained if it could be shown that the meaning of *gan* became (more) grammatical in nature through semantic bleaching, while this verb was undergoing phonological reduction, in addition to being found with the plain infinitive through reanalysis and having a function in a verse text. More specifically, I would argue that after the initial reanalysis of OE *onginnan* and OE and eME *aginnan* with a verb in the plain infinitive, it is possible to observe further reanalysis of these verbs into *gan* and *can* with a verb in the plain infinitive in the NME and ESc data sets (cf. Section 4.4.2.3.1). It is at this point that phonological reduction is taking place; instead of

affecting compound verbs in general, this time, the outcome of this process is more localised and affecting both *gan* and *can*, rather than *gan* only during a change from OE *onginnan* to OE and eME *aginnan*. As a result of this phonological reduction, the single morphemes *gan* and *can* emerge in NME and ESc, as in (59) below and (60) below.

- (59) *Pe slogth he fand þat him gan wiss*
 The path he found that him did mark
Tilward þe zate of paradis;
 Toward the gate of paradise;
 (1350-1420 *Cursor Mundi* p. 82; *HCET*)

‘He found the path that marked the way for him towards the gate of paradise;’

- (60) *Sancte Julyane than and his wyf*
 Saint Julian then and his wife
To God led thai sa thankful lyfe –
 To God led they so thankful life –
That thare dedis sa wele can stere
 That their deeds so well did guide
The lytil tyme that tha lifit here –
 The little time that they lived here –
 (1375-1420 *The Life of St. Julian Hospitaller* l. 474-476; *TEAMS*)

‘Saint Julian and his wife then led a life pleasing to God, who guided their actions well during the short time they lived here’

Overall, therefore, even though some sources argue that *gan* and *can* derive from *beginnen* (*OED* s.vv. *gin* v.¹; *begin* v.¹, Macafee 1992/93: §8.13), it is clear that the former two are not variants of the latter. This is because it is possible to see phonological differentiation between these three verbs in the NME and ESc data sets: both *gan* and *can* undergo phonological reduction as part of the process of grammaticalization, whereas *beginnen*, formed by the coalescence of OE prefix *be* ‘around’ (used to form a number of derivatives in OE) and the root *ginnan*, does not (*OED* s.v. *begin* v.¹).

3.5 Summary

In this chapter, I showed that the category of the verb undergoes a number of changes in the transition from OE to ME: a) the attrition of inflectional morphology; b) the reduction of vowel grades per verb; and c) the mixing of forms from other classes in the conjugation of the given verb, and transfer of verbs, either wholly or partly, into the weak conjugation. I also showed that *gan* and *can* in auxiliary use occur in the preterite only in ME and Scots, although this verb and its variant are still characterised by inflectional morphology in ME, especially in relation to the singular and plural dichotomy. The findings of my investigation also demonstrated that *gan* and *can* occur in the NME and ESc data sets without inflectional endings, with root vowel variation to form the preterite being orthographic. These kinds of changes on this verb and its variant are to be expected, as the process of attrition of inflectional morphology is more advanced in the northern dialects of ME and in Scots. The occurrence of this verb and its variant in the preterite only when in auxiliary use in the NME and ESc data sets could be contrasted with the presence of grammatical forms of *beginnen* in both finite and non-finite uses in these data sets. Taking this into account, I argued in this chapter that *gan* and *can* are fixed in the third person singular form in the preterite, this being an instance of decategorialization, as part of the process of grammaticalization. Furthermore, I also demonstrated that there is evidence of *gan* and *can* undergoing phonological reduction – this verb and its variant can be seen as single morphemes, following further reanalysis of this verb and its variant with the plain infinitive. However, I also argued that before this localised phonological reduction, the form *gan* arises out of phonological reduction of the *on* and *a* prefixes of OE *onginnan* and OE and eME *aginnan*. Nevertheless, I showed that the grammaticalization of these prefixes affected compound verbs in general, rather than *onginnan* and *aginnan* specifically. Overall, therefore, the verb *gan* and its variant *can* in the NME and ESc data sets appear to be morphologically and phonologically ‘lighter’, although only in terms of this they could be seen as satisfying one of the conditions in the process of grammaticalization, “[w]hen a form undergoes grammaticalization from a lexical to a grammatical form ... it tends to lose the morphological and syntactic properties that would identify it as a full member of a major grammatical category such as noun or verb” (Hopper & Traugott

[1993] 2003: 107).

Chapter Four – Changes in Complementation

4.1 Introduction

In this chapter, I investigate the complementation of *gan* and *can*, as well as that of *beginnen* in the NME and ESc data sets. In OE, both a plain infinitive (with an *-an* ending) and a *to*-infinitive were used (Warner 1993: 136). During the ME period, the *to*-infinitive became more frequently found with lexical verbs (cf. also Mossé 1952: 99-101, Mustanoja 1960: 514-515, Visser 1963/73: 948, Lightfoot 1979: 190, Fischer 1992: 317, Macafee 1992/93: §8.4, Denison 1993: 213-214, Burrow & Turville-Petre [1992] 1996: 48-49, King 1997: 179-180 and Moessner 1997: 150-152 amongst others), whereas one of the characteristics of some verbs classed as auxiliary verbs was their continued restriction to the plain infinitive (Warner 1993: 103, Fischer & van der Wurff 2006: 149). However, even though an auxiliary status have been argued for OE *onginnan* and OE and eME *anginnan* (Mustanoja 1960: 610-611, Visser 1969: 1580, Brinton 1988: 159-161, Sims 2008; 2014), but also OE *onginnan* and OE *beginnan* with the plain infinitive in Ælfric (Los 2000; 2005: 88-99), there was a competition between the plain infinitive and the *to*-infinitive that can be observed in relation to “the *ginnan*-verbs” in OE, but the use of the two infinitives with such verbs was not in free variation or random (Los 2005: 88, 97). This type of variation in relation to the choice of infinitive could also be found with *gan* and *can* in ME and Scots, although to a varying degree, with only *couth* being exclusively found with the plain infinitive (*DSL* s.vv. *can* v. 2; *gan* p.t., *OED* s.vv. *can* v.² I-II; *gin* v.¹, Visser 1969: 1579-1580, Brinton 1981: 156, Warner 1993: 139). Relying on the view that the *to*-infinitive was gradually used with verbs during the ME period is, therefore, insufficient in the case of *gan* and *can*, as well as *couth*, in ME and Scots, unlike in the case of *beginnen* in these varieties.

I begin by providing an account of changes in relation to the infinitive during the transition from OE to ME (Section 4.2). In Section 4.3, I provide an overview of different types of complementation found with *gan* and *can* in ME and Scots, based on the findings by Beschorner (1920), Mustanoja (1960), Visser (1969), Brinton (1981; 1988), King (1997), Macafee (1992/93), Ogura (1997) and Sims (2008; 2014)

amongst others. In this section, I also include my findings on the types of complementation found with *gan* and *can*, as well as *beginnen* in the NME and ESc data sets. I show that the former two almost exclusively occur with the plain infinitive in the NME data set and solely with this type of infinitive in the ESc data set, whereas the latter one is found with the (*for*) *to*-infinitive, the *at*-infinitive, but also the plain infinitive, with this choice of infinitive differing in the data sets and/or periods. In Section 4.4, I present a range of literature on the subject which attempts to account for the choice of infinitive found with the above-mentioned verbs (this will include references to the predecessors of *gan* and *can*, i.e. OE *onginnan* and OE and eME *aginnan*, as well as the predecessor of *beginnen*, i.e. OE *beginnan*). I also argue in this section that the occurrence of *gan* and *can* with punctual (achievement) verbs in the plain infinitive is the result of various reanalyses and analogies, with both *couth* as a variant form of *can* and *begouth* as a variant form of *beginnen* arising out of analogy. This chapter concludes that in the case of *gan* and *can* in the NME and ESc data sets, both reanalysis and analogy are part of the process of grammaticalization, in view of claims in the literature that the more grammaticalized variant takes the plain infinitive, whereas in the case of *beginnen*, it is possible to observe the use of this verb with the *to*-infinitive, as part of global changes affecting the infinitive in the transition from OE to ME.

4.2 Changes in the Infinitive Found with Verbs in the Transition from Old English to Middle English

While both the plain infinitive and the *to*-infinitive were used in OE (Warner 1993: 136), only some verbs were eventually restricted to the use of the former. This group of verbs were typically characterised by functional uses, although to a varying degree. Referring to both OE and ME, Fischer (1992: 317-318), for example, highlights the fact that the more grammaticalized an item, the more likely it is found with the plain infinitive, “it seems clear the more grammaticalised the matrix verb is, or, in other words, the emptier it is of referential meaning, the more likely it is that the [plain] infinitive is found”. Such verbs found with the plain infinitive constituted a restricted group. For instance, Fischer & van der Wurff (2006: 149) maintain that in OE, “[t]his class comprised the modals, verbs of physical perception (‘see’, ‘hear’)

and causatives (OE *lætan*, *biddan*, *hattan*).” Therefore, with their restriction to the plain infinitive, this group of verbs gradually formed a new class, namely that of auxiliary verbs, including modal (auxiliary) verbs. For example, Both Warner (2003: 136-139) and Roberts & Roussou (2003: 39) argue that irrespective of other complements found with these verbs at earlier stages, modal (auxiliary) verbs always took an infinitival complement, although rarely the *to*-infinitive (with the exception of *agan/ought*). Lexical verbs, on the other hand, were gradually found with the newly grammaticalized *to*-infinitive.

In its origins, a directional adverb or preposition, *to* initially began to function as an indicator of a purpose, but had lost this function by the eME period, eventually becoming an infinitival marker (Warner 1993: 136, Los 2005: 155-171, Fulk 2012: 107-108). In a similar way, *for-to* came to function as an infinitival marker, in the late 13th and the 14th centuries even alternating with the *to*-infinitive, especially in verse, where the function of both was that of a metrical device (Fischer 1992: 317). The use of the *to*-infinitive during in OE and ME was due to the fact that this infinitive started to replace *that*-clauses (Los 2005: 85-88, Fischer & van der Wurff 2006: 149), rather than being in competition with the plain infinitive and so occurring at its expense, as is traditionally assumed by Callaway (1913: 70), Bock (1931) and Sweet ([1898] 2014).⁶⁷ Given that there was no competition as such between the plain infinitive and the *to*-infinitive in the transition from OE to ME, the distribution of both types of infinitive was, therefore, not haphazard in ME. Quoting studies by Kaartinen & Mustanoja (1958) and Quirk & Svartvik (1970), Fischer (1992: 317), for example, argues that, at least, two parameters seemed to be of importance in determining the choice of infinitive with verbs: a) the intimacy of the verb and its infinitive; and b) the physical distance between the verb and its infinitive, both of these parameters appearing to be associated with the degree of grammaticalization. In other words, the more grammaticalized variant takes the plain infinitive.

⁶⁷ There are still grounds to assume that this kind of competition between the two structures in early OE when the *to*-infinitive had the function of a purpose adjunct (Fischer 1992: 317-318, Warner 1993: 137, Roberts & Roussou 2003: 39, Los 2005: 82-88, 99, Fischer & van der Wurff 2006: 149).

However, although there are grounds to assume that the competition between the two types of infinitive was in early OE only, when the *to*-infinitive had the function of a purpose adjunct (Fischer 1992: 317-318, Warner 1993: 137, Roberts & Roussou 2003: 39, Los 2005: 82-88, 99, Fischer & van der Wurff 2006: 149), in late OE (IOE) and, subsequently, in ME, too, a number of verbs continued to take the plain infinitive and the *to*-infinitive. According to Fischer & van der Wurff (2006: 149), such verbs included, “*þencan* ‘think’ in OE, and in ME also *helpen*, *maken*)” (cf. Fischer 1992: 317-318, Warner 1993: 137, Roberts & Roussou 2003: 39, Los 2005: 82-88). More importantly for this study, there was also a group of “monotransitive subject control verbs” in OE, namely “the *ginnan*-verbs”, where it would be possible to see some sort of competition between the plain infinitive and the *to*-infinitive, although “the two infinitives are not in free variation after the *ginnan*-verbs, and that their distribution is not random” (Los 2005: 88, 97). In ME and Scots, the variation in relation to the choice of infinitive was still found with the verb *gan*, but less so with the variant *can*, whereas the variant *couth* was recorded only with the plain infinitive in these varieties (*DSL* s.vv. *can* v. 2; *gan* p.t., *OED* s.vv. *can* v.; *gin* v.¹, Visser 1969: 1579-1580, Brinton 1981: 156, Warner 1993: 139, Moessner 1997: 114). The verb *beginnen*, on the other hand, occurred with the newly grammaticalized *to*-infinitive in ME and Scots. The types of complementation found with *gan* and *can*, as well as *beginnen* will be discussed in the following section.

4.3 Types of Complements Found with *Gan*, *Can* and *Beginnen* in Middle English and Scots

4.3.1 *Gan* and *Can*

In ME, *gan* mainly occurs with the plain infinitive, as in (1) below, although instances of this verb are also found with other types of infinitive, such as the *to*-infinitive, as in (2) below, and, even rarer, the *for-to* infinitive, as in (3) below.⁶⁸ In Scots, like in ME, this verb occurs with the plain infinitive, but in contrast to ME, instances with the *to*-infinitive are only recorded from 1500 onwards (*DSL* s.v. *gan*

⁶⁸ This instance of *gan* with the *for-to* infinitive is also provided in the *Corpus of Middle English Prose and Verse*, where it is taken from the Cotton Vespasian A.3 MS (cf. Göttingen MS Theol. 107, Trinity College, Cambridge MS R.3.8). However, in Oxford, Bodleian Library, Fairfax MS 14, <gan> is replaced with <began>, but still found with the *for-to* infinitive.

p.t., *MED* s.v. *ginnen* v1. 3b, *OED* s.v. *gin* v.¹ 1a, Beschorner 1920: 9-11, Mustanoja 1960: 531, Visser 1969: 1373-1375, 1379-1380, 1571-1576, Brinton 1981: 156; 1988: 120, Ogura 1997: 409, Sims 2008: 148-150). As *can* is a variant of *gan*, it is, therefore, not surprising to find that it also occurs with the plain infinitive in both varieties, as in (4) below and the *to*-infinitive, as in (5) below. (Although the *OED* s.v. *can* v.² I 2 lists *can* with the *to*-infinitive under the sense of ‘[b]egan, proceeded to do something’, I would argue that a non-ingressive meaning for this variant is also possible here, given that the verb *cast* ‘throw’ can be seen as a punctual (achievement) verb here, too as opposed to a semelfactive one). However, with reference to ME, both the *MED* (s.v. *can* v.) and the *OED* (s.v. *can* v.² I 2) have a considerably small number of instances of this variant with the *to*-infinitive than with the plain infinitive – both of these sources list one instance each of *can* with the *to*-infinitive from *Cursor Mundi*. Nevertheless, in Scots, the *DSL* (s.v. *can* v. 2a) shows that in this variety, only the plain infinitive is used with instances of this variant in this variety, whereas the *OED* (s.v. *can* v.² I 3) has one instance of the *to*-infinitive with *can* from Blind Harry’s *Wallace*.

- (1) *Fyue wyntyr wyþ hym gan y wone*
 Five winter with him did I live
 (a1400 (c1303) Robert Mannyng of Brunne *Handling Sin* 72; *MED* s.v. *ginnen* v. 3ba)
 ‘I lived with him for five years’
- (2) *þe belles of þe citee gonne to ryngge by*
 The bells of the city did/began to ring by
hem self ([L sonuerunt; *Higd.*(2): *did ryngge*)
 themselves
 (a1387 John Trevisa *Higden’s Polychronicon* 6.203; *MED* s.v. *ginnen* v. 3ba)
 ‘The bells of the city did/began to ring by themselves’
- (3) *þan gan a maister for to sai*
 then began a master for to say
 (a1400 *Cursor Mundi* 12168; Ogura 1997: 420; translation in the original)
 ‘then a master began to say’

- (4) *Whan Iudas þe traytour can þe mete*
 When Judas the traitor did you meet
 (a1425 (a1400) *Ihesu þat hast* 10; *MED* s.v. *can* v. a)

‘When Judas the traitor met you’

- (5) *And so on ane hys eyne he can [1648 began]*
 And so on one his eyes he did/began
to cast
 to cast
 (c1470 Henry *Wallace* iv.98; *OED* s.v. *can* v.² I 2)

‘And so, he did/began to pay attention to one [i.e. a woman]’

Crucially, however, whether the plain infinitive or the *to*-infinitive is used with *gan* and *can* would appear to depend on the degree of grammaticalization of this verb and its variant. For example, Moessner (1997: 114) argues about *gan* in OSc, making a comparison between it and modal (auxiliary) verbs, “[a]n expansion by *gin* (present) or *gan* (past) can be followed by the bare infinitive or by the *to*-infinitive. Although formally similar to the expansion by a modal auxiliary, it is syntactically and semantically different; syntactically, because it does not combine with other expansions, semantically, because *gin/gan* does not add a modal component” (cf. Warner 1993: 139), whereas the *OED* (s.v. *can* v.² I 1-2) gives more detail about the role of the infinitive found with the variant *can*, “[f]ollowed by bare infinitive, as a periphrastic auxiliary of the past tense”, but “[f]ollowed by *to*-infinitive [when used in the sense of] Began, proceeded *to* do something”. However, while *gan* and *can* permit variation in relation to the choice of infinitive, the variant *couth* shows no variation in relation to its complement, occurring only with the plain infinitive, in which case it functions as “a periphrastic auxiliary of the past tense” in OSc (*OED* s.v. *can* v.² II 4, cf. *DSL* s.v. *can* v. 2b, Visser 1969: 1579-1580). Overall, *gan* and *can* can occur with different types of infinitive, but whether this verb and its variant are used with the plain infinitive or not seems to reflect the degree of auxiliatation of this verb and its variant, with *couth* being only used with the plain infinitive, and so perhaps demonstrating a greater degree of grammaticalization in ME and Scots than *gan* and *can* do in these varieties.

However, some sources also report that *gan* and *can* are used with a form in *ende* or *ing* in ME, although no references to this use are made in NME and Scots. For example, Visser (1969: 1379) lists such instances, as in (6) below and (7) below, but also points out their scarcity. Nevertheless, there is doubt whether or not Visser's (1969) instances show use of the present participle, with Visser (1969: 1889, 1888-1890) pointing out that not in all cases, a form in *ende* is "equally convincing" when it is found with *gan* and *can*. I would argue that in relation to (6) below, while *endende* in this example could be in a form in <ende> (present participle) of the verb *end* (*MED* s.v. *enden* v.), its meaning could also be similar to the adverb *anende* 'in the end, finally' (*MED* s.v. *anende* adv.), with the translation of *the Epistle to the Hebrews* (8,8) in the *King James Version* of *The Bible* being 'Behold, the days come ... when I will make a new covenant'. With reference to (7) below, on the other hand, *barkyng* appears to be a variant form of the verb *berken* 'of a dog or fox: to bark; also, make a sound resembling a bark' (*MED* s.v. *berken* v.), despite the existence of the etymologically related *berking* functioning as a gerund with the meaning of 'the barking of a dog or a fox' (*MED* s.v. *berking* ger.). Other sources also raise doubt over instances of *gan* with a form in *ende* or *ing*, although their arguments are not entirely convincing. For example, Sims (2008: 150) maintains that it is only in relation to Visser's (1969) examples in *ing* that it is unclear whether we are dealing with instances of the present participle or a gerund, "[t]here is no question about *-ende* (the present participle affix), but *-ing* can be ambiguous between a participle and a gerund". However, De Smet (2013: 134) claims that although gerunds with verbal features sporadically begin to appear in IME, it was not until the ModE period that "the process of verbalization came into full swing and verbal gerunds became a truly indispensable part of English grammar."

- (6) *Lo, þe dayes schal come ... I endende schal*
 Lo, the days shall come ... I ending/finally shall
 gyne (L. consummabo) a newe testament
 establish a new testament
 (c144? Pauline *Epist.*, *Hebr.* 8, 8; Visser 1969: 1379)

‘Lo, the days shall come ... I will finally establish a new testament’

- (7) *He can barkyng as a foxe*
 He can bark as a fox

(a1500 MS Harley 1002; Visser 1969: 1379)

‘He can bark like a fox’

In addition to Visser’s (1969) doubtful instances of *gan* and *can* with complements in *end* and *ing*, there are also questions about the meaning of this verb and its variant in such constructions. For example, both Visser (1969) and Sims (2008) treat <gyne> in (6) above as ‘a weak auxiliary used with infinitives to form phrases denoting actions or events as occurring (rather than as beginning to occur): do, did’ (*MED* s.v. *ginnen* v. 3b). However, the meaning of this verb in this example is de facto different, listed in the *MED* (s.v. *ginnen* v. 2c) under the sense of ‘fig. establish (a new testament)’, while in the *King James Version* of *The Bible*, as referred to above, it has a causative meaning. Likewise, while the meaning of *can* in (7) above would probably be interpreted by Visser (1969) and Sims (2008) as similar to the meaning of *gan*, it would appear that *can* functions here as a modal (auxiliary) verb. The claim that the function of *can* is that of a modal (auxiliary) verb here can be further supported with information from the subsequent lines, as in (8) below, where the variant form *barken* is also found with the modal (auxiliary) verb *can* (I have highlighted both variants in this example). Overall, it appears that *gan* and *can*, as well as *couth*, only occur with the infinitive in ME and Scots, although the choice of infinitive seems to be different in terms of the verb and variety, but also, more importantly, whether or not the verb *gan* and its variant *can*, including the variant *couth*, are in auxiliary uses. Therefore, those instances of *gan* and *can* reported in Visser (1969) to take a form in *ende* or *ing* as their complements appear not to be instances of *can* with the present participle, as outlined above.

- (8)
- | | | | | | | | | | |
|-----------|-------------------|-----------------------|-------------|----------------|--------------|------------|------------|----------------|--|
| <i>At</i> | <i>my</i> | <i>house</i> | <i>I</i> | <i>have</i> | <i>a</i> | <i>jay</i> | | | |
| <i>At</i> | <i>my</i> | <i>house</i> | <i>I</i> | <i>have</i> | <i>a</i> | <i>jay</i> | | | |
| <i>He</i> | <i>can</i> | <i>make</i> | <i>mony</i> | <i>diverse</i> | <i>leye</i> | | | | |
| <i>He</i> | <i>can</i> | <i>make</i> | <i>many</i> | <i>diverse</i> | <i>lay</i> | | | | |
| <i>He</i> | <u><i>can</i></u> | <u><i>barking</i></u> | <i>as</i> | <i>a</i> | <i>fox</i> | | | | |
| <i>He</i> | <i>can</i> | <i>bark</i> | <i>as</i> | <i>a</i> | <i>fox</i> | | | | |
| <i>He</i> | <i>can</i> | <i>lowe</i> | <i>as</i> | <i>a</i> | <i>noxe</i> | | | | |
| <i>He</i> | <i>can</i> | <i>low</i> | <i>as</i> | <i>a(n)</i> | <i>ox</i> | | | | |
| <i>He</i> | <i>can</i> | <i>crecu</i> | <i>as</i> | <i>a</i> | <i>gos</i> | | | | |
| <i>He</i> | <i>can</i> | <i>cackle</i> | <i>as</i> | <i>a</i> | <i>goose</i> | | | | |
| <i>He</i> | <i>can</i> | <i>remy</i> | <i>as</i> | <i>a</i> | <i>nasse</i> | <i>in</i> | <i>his</i> | <i>cracche</i> | |
| <i>He</i> | <i>can</i> | <i>bray</i> | <i>as</i> | <i>a[n]</i> | <i>ass</i> | <i>in</i> | <i>his</i> | <i>cratch</i> | |

He can croden as a froge
He can croak as a frog
He can barken as a dogg
He can bark as a dog
He can cheteron as a wrenne
He can chatter as a wren
He can chateryn as a henne
He can cackle as a hen
He canne neye as a stede;
He can neigh as a steed;
Such a byrde yt were wode to fede
Such a bird it were wood to feed
 (MS Harley 1002, f. 72 recto; Orme 1995: 81)

‘At my house I have a jay / He can sing many a diverse song / He can bark like a fox / He can low like an ox / He can cackle like a goose / He can cry like an ass in its rack / He can croak like a frog / He can bark like a dog / He can chatter like a wren / He can cackle like a hen / He can neigh like a stallion; / It would be insane to feed such a bird!’

Let us now investigate the choice of infinitive with which *gan* and *can* are found in the NME and ESc data sets. The findings of this study, conveniently summarised in Table 39 and Table 40, show that in these data sets, this verb and its variant occur with verbs in infinitive only, as in (9) below, (10) below and (11) below.

Table 39

Infinitival Complements with gan and can in the NME Data Set in Period I and Period II

	<i>gan</i>				<i>can</i>	
	Plain-Inf		<i>To</i> -Inf		Plain-Inf	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Period I	17	100	-	-	8	100
Period II	74	99	1	1	12	100

Notes: *N* – Observed absolute frequency
 % – Per-cent of the type of complement per period

Table 40

*Infinitival Complements with **can** in the ESc Data Set in Period I*

		can	
		Plain-Inf	
		<i>N</i>	%
Period I		8	100

Notes: *N* – Observed absolute frequency
% – Per-cent of the type of complement per period

- (9) *þe woman þan hir dedes gan rew*
The woman then her deeds did repent
And þus scho said vnto Jesu,
And thus she said unto Jesus,
(1350-1420 *The Northern Homily Cycle P II* p. 82; *HCET*)

‘The woman then repented her sins and said to Jesus in this way,’

- (10) *The lyones fonde thay in hir dene,*
The lioness found they in her vale,
A knave childe laye sowkand hir then
A male infant lay sucking her then
And gan with the lyones to playe.
And began with the lioness to play.
Umwhile the childe sowkede hir pappe,
Sometimes the child sucked her breast,
Umwhile gan thay kysse and clappe;
Afterwards did they kiss and embrace;
(1420-1500 *Octavian* l. 442-446; *TEAMS*)

‘They [the men] found the lioness in her vale, a male infant lay sucking her then and began to play with her. Sometimes the child sucked her breast, afterwards, they kissed and embraced;’

- (11) *And ilkan can til oper sai,*
And each did/began to to other say,
“Quat a sing es þis!” coth þai
“What a sign is this!” said they
(1350-1420 *Cursor Mundi* p. 1015; *HCET*)

‘And each did/began to say to the other, “What a sign is this!” they said’

However, before I proceed with an interpretation of my findings, I outline here how I treat instances of complementation found in co-ordination with *gan* and *can*, as well as *beginnen* in this study. In the case of the infinitive in OE as part of co-ordination, Los (2005: 157-160, 213-214) argues that the co-ordinated elements are two separate clauses, which involve the ellipsis of the subject and verb of the second clause under identity. Evidence for such co-ordinated clauses in OE can also be found in ModE, as in (12) below. Los (2005: 160) maintains that evidence for this kind of treatment of co-ordinated clauses also exists in the literature, but there is a significant difference in relation to the judgements in normative and descriptive grammars. For example, referring to ModE, Aaron (1993: 48-49), Dupré (1995: 346) and Barnet & Stubbs (1995: 129) advocate that conjoined clauses are identical in form and function, whereas Greenbaum & Quirk (1990: 278) propose that such constructions may be different in form, as long as they are similar in function. On the basis of the latter, Los (2005: 160) argues, “[a]s long as both conjuncts are each well-formed complements with respect to the matrix verb, that seems to be enough to make such a coordination work, both in OE and ‘naive’ [ModE]”. Therefore, what is co-ordinated in (12) below is a pair of clauses, with the second verb being deleted, but, more importantly, also evidence that, for example, the verb *like* subcategorises for the *to*-infinitive and the present participle in ModE.

- (12) [Clause: Shirley likes to play tennis] and [Clause: ~~Shirley likes~~ watching basketball] (Los 2005: 159)

Although exclusively used with the plain infinitive in Period I, the verb *gan* shows some minor variation in the type of infinitive it selects in Period II in the NME data set: it predominantly occurs with the plain infinitive (99%), but there is also one instance of this verb with the *to*-infinitive (1%), as in (10) above. My findings suggest that *gan* shows a greater degree of grammaticalization, in view of claims in the literature that the more grammaticalized variant takes the plain infinitive, than is implied by descriptions in the literature on the choice of complementation by this verb (*DSL* s.v. *gan* p.t., *MED* s.v. *ginnen* v. 3b, *OED* s.v. *gin* v.¹ 1a, Beschorner 1920: 9-11, Mustanoja 1960: 531, Visser 1969: 1379-1380, 1575, Brinton 1981: 156; 1988: 120, Ogura 1997: 409, Sims 2008: 148-150). Statistically speaking, it is

necessary to find out whether there is a correlation between the difference in the frequency of the complement found with **gan** in Period I and Period II in the NME data set. With the log-likelihood for the plain infinitive with this verb being 4.10 and the *to*-infinitive being 0.65, the difference between frequency scores over periods is statistically significant at the level of $p < 0.05$ in relation to the former type of infinitive (95th percentile, 5% level), but not statistically relevant at the level of $p < 0.05$ with reference to the latter one (95th percentile, 5% level). Given that the lower the log-likelihood value, the least significant the difference is between two frequency scores, we should, therefore, reject the null hypothesis on there being no correlation between the differences in frequency scores in the case of the plain infinitive with **gan** between Period I and Period II, but accept it for the *to*-infinitive with this verb.

Let us now turn to the complement of **can** in the NME and ESc data sets. Unlike **gan**, this variant is exclusively used with the plain infinitive in these data sets, even though three instances of this verb are listed by the *OED* (s.v. *can* v.² I 2) to occur with the *to*-infinitive, too. However, although the number of examples given in this source need not reflect the actual frequency of this variant in both varieties, other sources dealing with this variant usually imply that **can** as a variant of **gan** is very rarely used with the *to*-infinitive in ME and Scots (*DSL* s.v. *can* v. 2a, *MED* s.v. *can* v.). Therefore, the findings presented here suggest that like **gan** in the NME data set, by exclusively occurring with the plain infinitive, **can** in this data set and in the ESc data set shows a greater degree of grammaticalization. In terms of statistical analysis, with the null hypothesis being that there is no correlation between the differences in the distribution of the plain infinitive over periods, the difference in the frequency of the plain infinitive with **can** in the NME data set between periods is not statistically significant at the level of p being < 0.05 , with the log-likelihood being 1.38 (95th percentile, 5% level); however, the opposite is the case for the plain infinitive found with this variant in the ESc data set, where the difference in the frequency of the infinitive between periods is statistically significant at $p < 0.0001$ (log-likelihood is 43.40, 99.99th percentile, 0.01% level). Therefore, the probability of the results – i.e. the differences in the distribution of the complement of **can** between the two periods in the NME and the ESc data set – happening by chance is more than 5% for the

plain infinitive with this variant in the former data set and less than 0.01% for this kind of complement in the latter data set.

4.3.2 *Beginnen*

While *gan* and *can* are mainly or exclusively found with the plain infinitive in ME and Scots, the verb *beginnen* occurs with the plain infinitive in ME, as in (13) below, the *to*-infinitive, as in (14) below (including the *for-to* infinitive) (Visser 1969: 1373-1375, Brinton 1981: 181-182; 1988: 118-119, Ogura 1997, Sims 2008: 152), but also with the present participle/verbal noun (Brinton 1981: 182-183, Sims 2008: 152), with zero complement, as in (15) below, or other types of non-infinitival complements, such as an NP (Ogura 1997). In Scots, on the other hand, *beginnen* is found to occur with the *to*-infinitive, with the *DSL* (s.v. *begin* v. 1) maintaining that this verb is only “rarely [found] with the ellipse of *to*”. In relation to *begouth*, forms of this verb are recorded to occur exclusively with the *to*-infinitive⁶⁹ in the *DSL* (s.v. *begin* v. 1b b) and the *OED* (s.v. *begin* v.¹ 1a), as in (16) below, in addition to one instance of this verb functioning as an independent verb before 1500, mentioned by the *DSL* only, as in (17) below. These statements, therefore, suggest that while more lexical verbs were occurring with the *to*-infinitive during the transition from OE to ME, this global change was more advanced in Scots than in the other varieties of ME, although the above sources do not distinguish between NME and SME. Furthermore, the above observations on the complementation of *beginnen* in ME and Scots can be contrasted with those presented by Los (2000; 2005: 88-99), who shows that in Ælfric, beside OE *beginnan* in auxiliary use when found with the plain infinitive, there were also instances of OE *beginnan* with the *to*-infinitive.

- (13) *Ful yern on godd bi-gun þai call*
 Very earnestly on God began they call
 (a1400 (a1325) *Cursor Mundi* 5942; *OED* s.v. *begin* v.¹ 1a)

‘They very earnestly began to call on God’

- (14) *Þe day beganne to daw*
 The day began to dawn

⁶⁹ There is one instance of the variant *begouth* with the plain infinitive, but its occurrence is in Boece’s *The History and Chronicles of Scotland* dated 1531 and so, outwith the scope of this study.

(c1475 (?c1425) *Avowing of King Arthur* 469; *OED* s.v. *begin* v.¹ 1a)

‘The day began to dawn’

- (15) *Now þis prolouge wil we bliþ,*
Now this prologue will we cease,
In crist nam our bok begin
In Christ name our book begin
(a1300 *Cursor Mundi* 266; *OED* s.v. *begin* v.¹ 2a)

‘We will now cease this prologue, [and] begin our book in Christ’s name’

- (16) *Than in Scotlande the Scottys Begouthe to*
Then in Scotland the Scots began to
regne
reign

(c1420 Andrew of Wyntoun *The Original Chronicle of Scotland* IV 1103; *DSL* s.v. *begin* v. 1b b)

‘Then the Scots began to reign in Scotland’

- (17) *Quhen he begouth this doughtus werk*
When he began this doubtful work
(c1420 Andrew of Wyntoun *The Original Chronicle of Scotland* V 4750; *DSL* s.v. *begin* v. 2b b)

‘When he began this doubtful work’

There is a discrepancy in the available literature about the dating of *beginnen* occurring with the above types of complementation in ME, but the findings presented by the sources listed in this paragraph point to an increasing use of the *to*-infinitive with this verb overall. For example, Visser (1969: 1373) argues about the complements of *beginnen*, “[u]p to about the middle of the fifteenth century plain infinitive and *to*-infinitive occur side by side; afterwards only the *to*-infinitive is normally used ... [f]or *to* + infinitive can be attested from the beginning of the fourteenth century; after 1500 it is only sporadically met with”, whereas Brinton (1988: 118-119), who, like Visser (1969), is interested in the qualitative, rather than quantitative, data, maintains, “the plain infinitive is found occasionally, the prepositional infinitive is usual by late Middle English”, adding that the presence or absence of the *to* (or *for-to*) is “a matter of authorial preference or prosodic exigency”. Finally, Sims (2008: 153) argues, “while the bare infinitive does occur

with *beginnen* during the ME period, its use is extremely minimal compared to the *to*-infinitive". The observations by Visser (1969), Brinton (1988) and Sims (2008) summarised in this paragraph suggest that unlike in the case of *gan* and *can*, the verb *beginnen* is not characterised by the same degree of auxiliarity in ME as are this verb and its variant. Furthermore, the remarks made by these sources also imply that the verb *beginnen* participates in the global language change concerning the *to*-infinitive, whereby it is increasingly found with this type of complement, rather than the plain infinitive, by then restricted in use to a select number of verbs (Warner 1993: 103).

Instances of *beginnen* have also been found to occur with a form in *ende* or *ing*. Such constructions have been recorded in ME, but not in Scots (*DSL* s.v. *begin* v., Macafee 1992/93, King 1997). Visser (1973: 1890) elaborates on the use of such constructions, arguing that they were also found with either *a* or *an*, the former of which was characterised by the loss of the final consonant in fixed idiom to reflect an unstressed pronunciation of words in proclitic use, as in (18) below; both are variants of the preposition of *on*. However, the syntactic status of such constructions can be interpreted as either nominal or verbal, this highlighting the fuzzy categorisation of such formations. Smitherberg (2005: 37), for example, maintains, "[i]n cases where the *a*-form occurs with without either an *of*-phrase or a direct object (as in *I am a-singing*), there is thus only a weak indication that the *-ing* form is nominal; and when a direct object follows the *-ing* form (as in *I am a-singing a song*), there is a clear syntactic indication that the construction is verbal (the direct object), but only a vague indication that it is nominal (the reduced preposition)." Furthermore, I would also argue that an important property of constructions consisting of the prepositions *a* and *an* and a form in *ing* is that such constructions illustrate grammaticalized uses. Prepositions in such grammaticalized constructions are in such cases subject to extension in that they do not exemplify spatial uses in physical terms, but in metaphorical ones, indicating an action of engagement in some kind of activity, but also the process. However, as with other grammaticalized constructions, it is the syntactic mechanisms and processes involved in the formation of such constructions that make them grammaticalized, rather than meaning changes on their own.

- (18) *Quen þai vnderstode þis worde A soruing*
 When they understood this word A sorrowing
þai bigan
 they began
 (13.. *Cursor Mundi* 15518; Visser 1973: 1890)

‘When they understood what was said, they began crying out’

When the above observations about the complementation of *beginnen* are compared to my findings on the complementation found with this verb in the NME and ESc data sets, which are summarised in Table 41 and Table 42, the following patterns can be identified about this verb: a) *beginnen* occurs as an independent verb in both data sets and periods, where it takes zero complementation and other types of non-infinitival complements, as in (19) below and (20) below; b) *beginnen* also takes different infinitival complements, as in (21) below, (22) below, (23) below, although the choice of the infinitival complementation will typically differ in the data sets and periods; and c) *begouth* as a variant of *beginnen* only occurs as an independent verb in the ESc data set, even though its instances are mainly recorded with infinitive in Scots by the *DSL* (s.v. *begin* v. 1b b) and the *OED* (s.v. *begin* v.¹ 1a). Therefore, it is both a *to*-infinitive with which *beginnen* is found and the occurrence of this verb as an independent verb, in which case this verb occurs with zero complementation and other types of non-infinitival complements, that make this verb different in relation to its complementation from *gan* and *can* in the NME and ESc data sets.

Table 41

*Complements with **beginnen** in the NME Data Set in Period I and Period II*

		<i>beginnen</i>							
		Other		Plain-Inf		(For) to-Inf		At-Inf	
		<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Period I		24	47	1	2	25	49	1	2
Period II		38	22	-	-	128	74	7	4

Notes: *N* – Observed absolute frequency
 % – Per-cent of the type of complement per period

Table 42

Complements with *beginnen* in the ESc Data Set in Period I and Period II

	<i>beginnen</i>							
	Other		Plain-Inf		(For) to-Inf		At-Inf	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Period I	7	88	-	-	1	13	-	-
Period II	55	83	-	-	11	17	-	-

Notes: *N* – Observed absolute frequency
 % – Per-cent of the type of complement per period

(19) *Thurgh vertu of þe haly gaste,*
Through virtue of the holy ghost,
So þat he gude werkes bigin,
So that he good works begin,
And fulfils penance for his sin
And fulfils penance for his sin
 (1350-1420 *The Northern Homily Cycle* P III p. 137; *HCET*)

‘through [the] virtue of the Holy Ghost, so that he [the man] begins good works, and fulfils penance for his sin’

(20) *AN-OTHur LARD & LARGLY BEGWD QUHAR*
another lord & largely began where
HIS PREDECESSOR LEFTE ERYT &
his predecessor left ploughed &
LABORIT YE SAID LAND
laboured the said land
 (1459 *Arbroath Abbey*, “*Registrum Nigrum*” 341; *LAOS*)

‘another lord and largely began where his predecessor left ploughing and labouring the said land’

(21) *And if þai any tym begyn till erre, thurgh*
And if they any time begin to err, through
ignorance or frelte
ignorance or frailty

(1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 11; *ICoMEP*)

‘And if they anytime begin to err, through ignorance or frailty’

- (22) *And he was born in-to þer oratorie, &*
And he was brought into their oratory, &
þe monkis began att say þer psalters for
the monks began at say their psalters for
hym
him

(1420-1500 *Alphabet of Tales* p. 341; *ICoMEP*)

‘And he was brought into the oratory, and the monks began to recite their psalters for him’

- (23) *þan bigan þai cal and cri,*
Then began they call and cry,
þat godd o þam suld ha merci
That God on them should have mercy

(1350-1420 *Cursor Mundi* p. 424; *HCET*)

‘Then they began to call and cry to God to have mercy on them’

Of all instances of this verb in the NME and ESc data sets, the following observations can be made with regards to the choice of complementation of *beginnen*: instances of this verb with zero complementation and other types of non-infinitival complements constitute the majority in the ESc data set in Period I and Period II (88% and 83%, respectively), as opposed to instances of this verb with infinitival complements in the NME data set in Period I and Period II (53% and 78%, respectively). In terms of statistical analysis, the null hypothesis is that the difference in frequency of the verb *beginnen* as an independent verb and with infinitive shows no correlation between periods in the NME and ESc data sets. With the log-likelihood for this verb as an independent verb in the NME data set being 3.43 and the log-likelihood for this verb with infinitive being 11.2, the difference in frequency of this verb in the former function is not statistically significant over periods at the p level being < 0.05 (95th percentile, 5% level), but it is significant in the latter function at the $p < 0.001$ (99.9th percentile, 0.1% level). Therefore, the null hypothesis should be accepted for *beginnen* functioning as an independent verb in this data set, but rejected for this verb with infinitive. In relation to *beginnen* in the ESc data set, the

log-likelihood values are 1.81 when this verb functions as an independent verb and 0.05 when it is found with infinitive, the difference in frequency of this verb in both functions not being significant over periods at the p value being < 0.05 (95th percentile, 5% level) in both cases – this shows that there is no correlation between the differences in the frequency of different types of complementation of this verb over periods in the ESc data set.

With regards to infinitival complementation only, while the verb *gan* and its variant *can* almost exclusively occur with the plain infinitive in the NME and ESc data sets, the verb *beginnen* clearly shows a preference for the (*for*) *to*-infinitive in these data sets in Period I and Period II, as in Table 43 and Table 44. So, in the NME data set, out of all types of infinitive, *beginnen* mainly selects the (*for*) *to*-infinitive in Period I (93%) and Period II (95%), as in (21) above, with isolated instances of this verb found with the *at*-infinitive in both periods, too (4% and 5%, respectively), as in (22) above, in addition to one instance of this verb with the plain infinitive in Period I (4%), as in (23) above. The null hypothesis is that there is no correlation between the differences in frequency of occurrence of different types of complement found with *beginnen* over Period I and Period II in the NME data set. The following are the log-likelihood values for the types of infinitival complementation with *beginnen*: 2.56 for the plain infinitive, 11.17 for the (*for*) *to*-infinitive and 1.09 for the *at*-infinitive. Therefore, the difference in the frequency of the plain infinitive and the *to*-infinitive is not significant at the p level < 0.05 (95th percentile, 5% level), this meaning that we should accept the null hypothesis. However, the alternative hypothesis, namely that there is a correlation between the differences in the frequency of the (*for*) *to*-infinitive with *beginnen* should be accepted at the level of $p < 0.001$ (99.9th percentile, 0.1% level). In the ESc data set, on the other hand, out of all types of infinitive, *beginnen* exclusively occurs with the (*for*) *to*-infinitive in Period I and Period II. Statistically, with the log-likelihood for this type of infinitive with this verb being 0.05, there is no correlation between the differences in the distribution of the (*for*) *to*-infinitive at $p < 0.05$ over periods in the ESc data set (95th percentile, 5% level).

Table 43

*Infinitival Complements with **beginnen** in the NME Data Set in Period I and Period II*

	<i>beginnen</i>					
	Plain-Inf		(For) to-Inf		At-Inf	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Period I	1	4	25	93	1	4
Period II	-	-	128	95	7	5

Notes: *N* – Observed absolute frequency
% – Per-cent of the type of complement per period

Table 44

*Infinitival Complements with **beginnen** in the ESc Data Set in Period I and Period II*

	<i>beginnen</i>					
	Plain-Inf		(For) to-Inf		At-Inf	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Period I	-	-	1	100	-	-
Period II	-	-	11	100	-	-

Notes: *N* – Observed absolute frequency
% – Per-cent of the type of complement per period

Although the difference between the *to*-infinitive, the *for-to* infinitive and the *at*-infinitive is not as interesting as that between them and the plain infinitive, especially given that the occurrence of the plain infinitive with *gan* and *can* would suggest a greater degree of grammaticalization than of this verb and its variant than that of *beginnen*, there are still some interesting observations to make about the variation in relation to the (*for*) *to*-infinitive and the *at*-infinitive. The former found with *beginnen* in the NME and ESc data sets is realised as either <til(l)> or <to>, as in

(21) above, whereas the latter found with this verb in the NME data set is realised as <at(t)>, as in (22) above. The fact that the *at*-infinitive is only found in the NME data set appears to be in line with statements in the literature, where, for example, Macafee (1992/93: §8.4) states that this infinitive is found in Northern English, but in Scots in fossilised phrases like *ado* (= *at* + *do* ‘to do’) only. Macafee (1992/93: §8.4) and King (1997: 179-180) also argue that in Scots, the occurrence of <til(l)> and <to> is phonologically conditioned: it is realised as <til> before vowels and [h]. However, the use of either <til(l)> or <to> with the verb *beginnen* before vowels and [h] does not appear to be consistently applied in the NME and ESc data sets. For example, in the former data set, the *to*-infinitive is realised as either <til> or <to> before vowels and [h] and other consonants, whereas in the latter one, it is realised as <til> before a [h], but <to> before a vowel. Therefore, my findings show that the choice between <til> or <to> does not appear to be exclusively phonologically conditioned in the NME and ESc data sets, although, admittedly, there is not enough data in the ESc data set to fully ascertain the pattern. Furthermore, although the generalisation given by Macafee (1992/93) and King (1997) was specific to Scots, it would not necessarily be expected to apply to NME, and so, to the NME data set.

4.4 Accounting for Infinitival Complements of *Gan*, *Can* and *Beginnen* in Middle English and Scots

Since *gan* and *can* mainly occur with the plain infinitive in ME and Scots, while *beginnen* usually with other types of infinitive in these varieties, a number of different proposals have been made in the literature to account for the choice of infinitive of the above verbs, but also in relation to other verbs in general. These proposals have been based on metrical, stylistic and (semantic-) syntactic criteria. In what follows, I present an overview of the metrical and stylistic proposals, noting potentially grey areas that might prevent a satisfactory explanation as to why *gan* and *can* have a tendency to predominantly select the plain infinitive, whereas *beginnen* mainly opts for other types of infinitive, most commonly the *to*-infinitive and the *for-to* infinitive, in ME and Scots, including the NME and ESc data sets. However, while such proposals all have some merit in that they can perhaps be used to attempt to justify the use of any type of complementation with the above verbs, some of them

can also be perceived as subjective and so unsatisfactory in an effort to account for a specific type of infinitive with such verbs. Given this, I argue that only the syntactic proposal, where the final outcome is the result of reanalysis and analogy, can be successfully used in order to account for the choice of infinitive with which the above-mentioned verbs occur. Let us start with metrical and stylistic considerations.

4.4.1 Metrical and Stylistic Considerations

As *gan* and *can* are predominantly or exclusively used in verse in ME and/or Scots, Beschorner (1920: 11), Mustanoja (1960: 531), Kerkhof (1966: 154) and Brinton (1981: 156) propose that the choice of infinitive found with this verb and its variant, where references are made to *can* in these sources, depends on metre. For example, Mustanoja (1960: 531) argues that “[i]n poetry, the choice between the two forms of the infinitive [with *gan*] is often purely a matter of metre”, as in (24) below and (25) below. In a similar way, Kerkhof (1966: 155) maintains, “[a]lthough gan accompanied by an infinitive with (for) to often keeps its original meaning better than with the plain infinitive, the choice between the two is sometimes conditioned by metre”. A number of other studies also highlight the fact that this verb and its variant take both types of infinitive, but when the plain infinitive is chosen, the verb *gan* and its variant *can* is likely to have auxiliary uses, although this is not always the case. For example, Warner (1993: 139) points out that in terms of complementation, the verb *gan* in ME behaved differently from other verbs, especially modal (auxiliary) verbs, “*ginnen* may well show a preference for the plain infinitive, especially in weakened senses. But it occurs not infrequently with the *to*-infinitive, and it has yet to be demonstrated that any particular subset of its uses is restricted to the plain infinitive.” Overall, therefore, since the use of this verb and its variant is in verse, it would appear that there is a necessity to fit linguistic forms such as the plain infinitive and *to*-infinitive to metre, despite the fact that in auxiliary uses, *gan* and *can* show a preference for the plain infinitive in ME and Scots.

- (24)
- | | | | | | | | | |
|------------|---------------|------------|--------------|-----------------|---------------|------------|-----------|----------------|
| <i>Til</i> | <i>that</i> | <i>the</i> | <i>belle</i> | <i>of</i> | <i>laudes</i> | <i>gan</i> | <i>to</i> | <i>rynge</i> , |
| To | that | the | bell | of | Lauds | began | to | ring, |
| <i>And</i> | <i>freres</i> | <i>in</i> | <i>the</i> | <i>chauncel</i> | <i>gonne</i> | | | <i>synge</i> |
| And | friars | in | the | chancel | did/began | to | | sing |

(Chaucer *Canterbury Tales: The Miller's Tale* 3655-3656; Mustanoja 1960: 531)

‘To that, the bell of Lauds began to ring, and friars did/began to sing in the chancel’

(25) *Unto the wode I gan to fare*
Onto the wood I did/began to travel
(Gower *Confessio Amantis* I 110; Mustanoja 1960: 531)

‘I did/began to travel to the forest’

However, while metrical reasons can explain some of this kind of variation that is mentioned in the literature in relation to the verb *gan* and its variant *can* found with either the plain infinitive or the *to*-infinitive, the assumption should usually be that if a poet generally had a choice in what he did, and if he decided on a particular variant, he had a reason to do so beyond the demands of metre. Therefore, even though poets might have variation at their disposal between both types of infinitive with the verb *gan* and its variant *can* in ME and Scots, constructions that are ungrammatical would not be generally accepted. In view of this, the metrical approach would be disputable in relation to constructions permitting only one type of infinitive, such as the predecessors of the modal (auxiliary) verbs, which “were distinguished from most other verbs by their morphology, which is ‘preterite present’ [except for *wile*] and by the fact that an infinitive in construction with these verbs always lacked *to*” (Warner 1993: 94). What is more, although variation is often exploited to supplement the metre (cf. *ne* in Chaucer), this metrical proposal suggests that the choice between the plain infinitive and other types of infinitive is not determined by any linguistic factors, with only the number of syllables in a line of verse being taken into account whether or not a particular type of infinitive can or should be used. For the above reasons, I would, therefore, argue that metrical considerations alone are insufficient as a reliable diagnostic to determine the choice of infinitive with which *gan* and *can* will occur; this could also be extended to *beginnen* and the type of infinitival complementation found with it, especially that this verb is found with both the plain infinitive and, more often, other types of infinitive (Visser 1969: 1373-1375, Brinton 1981: 181-182; 1988: 118-119, Ogura 1997, Sims 2008: 152).

While the above sources argue that the choice of infinitive with *gan* and *can* would be determined by metre, Ogura (1997: 417) proposes that style can be an important factor in this kind of decision, “the choice of between Inf and *to*-Inf with [OE *onginnan* and *aginnan*, as well as *beginnan*] is not strict, and in the case of early Middle English *ginnen* and *beginnen* it seems stylistic rather than syntactic”. However, style is based on word choice and placement. Although this suggests that the above-mentioned verbs are used with both types of infinitive in eME, chosen on the basis of which of them sounds the best and so being subjective, Ogura (1997) does not show how style could play a particular role in the choice of infinitive with *gan* and *can* in ME. Instead, she maintains that “[i]t is not the meaning of the verb in the infinitive that can be decisive factor of the choice” and “[a] syntactic problem may occur when a pre-verb with *to* and a non-pre-verb were synonymous; it would be difficult to determine whether the construction that follows *-ginnan* (> *-ginnen*) [i.e. *gan*] is *to*-verb in the infinitive or a preposition *to* + a non-pre-verb in the infinitive” (Ogura 1997: 418), as in (26) below and (27) below. However, both *to-reosen* and *reosen* are separate lexemes, with the former one being a compound verb consisting of the OE derivational prefix *to* (*OED* s.v. *to-* prefix² 1-2), this prefix being etymologically different from the infinitival *to*. Furthermore, both verbs also have different senses, the former verb meaning ‘to fall to pieces, fall into ruins; to decay’, whereas the latter verb denoting the following: ‘to fall or fall down; to fall (in battle), perish; to descent, to drop’ (*OED* s.vv. *to-reose* v.; *reose* v.); the meanings of these verb are, therefore, not synonymous, but similar, as shown in (28) below; in this respect, they would pose a difficulty if indeed, a choice between both verbs would be a matter of style, not meaning, only.

- (26) A: *a niht he gon to-reosen*
 In night it [pret.] fall
 B: *a niht he gan to-rese*
 (Lazamon *Brut* 15483; Ogura 1997: 418; translation in the original)

‘in the night it fell’

- (27) & *ælche niht he gon reosen*
 and each night it [pret.] fall
 (Lazamon *Brut* 15493; Ogura 1997: 418; translation in the original)

‘and each night it fell’

(28) A *dæi* *heo* *leiden* *þene* *wal:*
A day they laid the wall:
a *niht* *he* *feol* *ouer* *al.*
a night it fell over all.
a *marwe* *heo* *hine* *arædden:*
the morrow they it reared:
a *niht* *he* *gon* *to-reofen.*
the night it did to-tumble.
fulle *feouen* *nihte:*
full seven night:
ælche *dæi* *heo* *hine* *arædden:*
each day they it deliver:
& *ælche* *niht* *he* *gon* *reofen*
& each night it did fall

(Lazamon *Brut* 15480-15487; Madden 1847: 224; translation in the original)

‘In *the* day they laid the wall, in *the* night ‘it’ fell over all; in *the* morrow they reared it, in *the* night it gan to tumble! Full *a* se’night [week] ‘so *it* [thus this work] them served; ‘each day they raised it, and each night it gan fall!’

Ogura (1997: 418) then seems to imply that there is a choice of infinitive found with the verb *gan* and its variant *can*, as well as the verb *beginnen*, in addition to other such etymologically related verbs; this choice appears to be much wider than that suggested in the literature, though, “*began*, *agan*, *gan/gon*, and *con/can* (and later, *did*) can be used alternatively, especially in poetry. They may take either Inf, *to*-Inf, *for to*-Inf, *till*-Inf, or *at*-Inf”. Given her proposals that these verbs select different types of complement on the basis of style, Ogura (1997) suggests that the choice of infinitive is in free variation (effectively being the result of stylistic preferences), considering the occurrence of the OE singular form *onginnan* in the preterite with “the infinitive”, i.e. without specifying the type of infinitive, as a sign of increasing grammaticalization (404). According to her, towards the ME period, both increasing grammaticalization and phonological reduction of OE *onginnan* and OE and eME *aginnan*, therefore, result in *gan* with infinitive in auxiliary use in ME. In her view, *gan* and *can*, as well as *beginnen* take different complements, in spite of claims in the literature that the more grammaticalized variant take the plain infinitive – Warner (1993: 139) argues that in ME, the group of verbs taking the plain infinitive, with the

shared subject, semantic included the descendants of the former preterite-present verbs in OE (*cann*, *dearr*, *mæg*, *mot*, *sceal*, *þearf*, *uton* and *wile*, but also *mun* and, at a later date, the periphrastic *do*), while the other verbs, which were characterised by subject raising semantics, were initially categorised for both the plain infinitive and the *to*-infinitive, with the latter ultimately becoming the norm (verbs of sense perception take the plain infinitive, though, and a significant number of verbs of “ordering, permitting or causing”). Overall, therefore, while Ogura (1997) suggests that the choice of infinitive in such cases is determined by stylistic factors, it is the choice of infinitive that reflects the degree of grammaticalization of a construction.

4.4.2 Reanalysis and Analogy-Based Syntactic Approach

With the metrical and stylistic approaches being difficult to prove, sources like Los (2000; 2005: 82-99) and Sims (2008) amongst others use proposals based on semantics and, more specifically, syntax, in particular reanalysis and analogy, in order to account for the preference of OE *onginnan*, i.e. the predecessor of *gan* and *can*, for the plain infinitive, whereas that of OE *beginnan*, i.e. the predecessor of *beginnen*, for the *to*-infinitive. The role of reanalysis⁷⁰ as the primary mechanism behind grammaticalization leading to auxiliatation has been acknowledged in the literature (cf. also Traugott 1972, Langacker 1977, Lightfoot 1979, Heine & Reh 1984, Anttila 1989 and Warner 1993 amongst others), although Haspelmath (1998), for example, argues that there is no relationship between reanalysis and grammaticalization. Heine & Kuteva (2002: 5), however, maintain that whether or not grammaticalization involves reanalysis is “a theory-dependent issue”, with a number of sources (Langacker 1977, Heine & Reh 1984, Harris & L. Campbell 1995: 61-96, Haspelmath 1998, Newmeyer 1998: 241-251) providing different definitions of this mechanism of change. Other sources dealing with different aspects of reanalysis include Ard (1976), Lord (1976), Haiman (1977), Timberlake (1977), Zubizarreta (1985), Heine, Claudi & Hünemeyer (1991a), Hock (1991), Hopper (1992), Rizzi (1992), Harris & L. Campbell (1995), Gaeta (1998), Jones & Singh (2005) and Heine & Kuteva (2005) amongst others. For discussion on treatment of

⁷⁰ The term ‘reanalysis’ is also found under different names in the literature: reinterpretation, rebracketing, relabelling and restructuring.

reanalysis in the literature, cf. L. Campbell (2001: 141-142). One of the most comprehensive descriptions of reanalysis can be found in Langacker (1977), who defines it as, “change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its structure manifestation ... [it] may lead to changes at the surface level ... but these surface changes can be viewed as the natural and expected result of functionally prior modifications in rules and underlying interpretation” (57-58).

Although the effects of reanalysis can eventually be observed on the surface level, this mechanism of change primarily affects underlying structure, i.e. it involves a change in structure. Langacker (1977: 62) argues, “any aspect of morphological, syntactic, or semantic structure, i.e. anything more abstract than the surface level ... These structural features include the location of morpheme and clitic boundaries, surface constituent structure, underlying constituent structure, the semantic value or syntactic function of morphemes, and so on.”⁷¹ The mechanism of reanalysis can affect the underlying structure in a number of ways. Harris & L. Campbell (1995: 61) specify, “[r]eanalysis directly changes underlying structure, which we understand to include information regarding at least (i) constituency, (ii) hierarchical structure, (iii) category labels, (iv) grammatical relations, and (v) cohesion”. Hopper & Traugott ([1993] 2003: 51) add that a single instance of reanalysis can often involve a number of these characteristics, occurring not in one single step, but being the result of a set of smaller changes. Traugott & Trousdale (2013: 74) refer to this kind of change as

⁷¹ Langacker (1977: 64) identifies two types of reanalysis: “resegmentation” and “syntactic/semantic reformulation”: the former refers to “the more superficial side in syntactic structure ... [i.e.] the occurrence and placement of morpheme boundaries”, whereas the latter pertains to “more abstract aspects of semantic and syntactic structure.” He continues that both types are independent, where one may occur without the other, but he states that frequently, both resegmentation and reformulation are involved in the process of reanalysis. Within resegmentation, Langacker (1977: 64-65) states that there are “‘boundary loss’, ‘boundary creation’, and ‘boundary shift’”. All three kinds of resegmentation can occur, but to a varying frequency: a) boundary loss, which involves “the downgrading of a boundary from a clitic to a morpheme boundary”, is the most frequent; b) boundary creation is “probably the least common”; and c) boundary shift is “apparently intermediate” in frequency. Nevertheless, he admits himself that “[t]hese estimates are naturally very crude”. Syntactic/semantic reformulation, according to Langacker (1977: 79), refers to “aspects of structure more abstract than the occurrence and placement of morpheme boundaries”; these aspects consist of rules, semantic and syntactic categories as well as configurations. Furthermore, Langacker (1977: 79) adds that reformulation may involve resegmentation, but it cannot be considered as a subcategory of resegmentation. Finally, as with resegmentation, three types of reformulation exist: a) the loss of semantic elements; b) their addition; and c) a shift in their value.

‘gradualness’ (Lichtenberk 1991), i.e. “a phenomenon of change, specifically discrete structural micro-changes and tiny-step transmission across the linguistic system”. An example of reanalysis incorporating a number of the above characteristics is the ModE *be going to* construction, which consists of changes in constituency structure and reassignment of morphemes into a different semantic-syntactic category from the coalesce of *be*, the main verb, progressive aspect and the preposition of purpose to a tense marker, ‘expressing a plan or intention that something will happen (usually soon), or making a prediction that something will happen, based on present events or circumstances’ (*OED* s.v. *go* v. 51).

The story of the verb *gan* and its variant *can* in ME and Scots can only be associated with reanalysis after it is put in the larger context of the rise of auxiliary verbs as a new category, separate from that of lexical verbs. This, in turn, is part of that key feature of the history of English, which not only involved the change from a synthetic language to an analytical one (word order changes), but also a series of instances of grammaticalization, resulting in an emergence of functional items. Such free form functional items, i.e. prepositions, articles, auxiliary verbs, including modal (auxiliary) verbs etc., were employed from the lexicon to perform grammatical functions during the transition from OE to ME. Below I discuss the role of reanalysis in the rise of auxiliary verbs.

4.4.2.1 *Rise of Auxiliary Verbs*

In OE, all verbs, including the predecessors of ModE auxiliary verbs, shared some characteristics: in negation, they were preceded with the negative particle *ne*, whereas in questions, they inverted to clause-initial position. As already referred to in Chapter Three, some verbs that were to form the class of auxiliary verbs were becoming morphologically distinct in OE: in addition to having preterite-present morphology, whereby the present tense had been formed with the morphology of the preterite (cf. Section 3.2), such verbs were characterised by the negative fusing with forms of several of them, i.e. *ne wille* ‘not intend’ > *nille* and *ne wæs* ‘not was’ > *næs*. Furthermore, the predecessors of ModE modal (auxiliary) verbs were also semantically different: they denoted completed action with the effect on the present

state (“‘have come to be X’”) (Hopper & Traugott [1993] 2003: 56). Apart from the above-mentioned morphological and semantic aspects (including word order changes), during the ME period, further changes occurred, leading to the emergence of the new class of auxiliary verbs. These changes include: the development of a new negative particle *not*, found after the verb, and the use of preterite forms of the predecessors of modal (auxiliary) verbs *would*, *might*, *could* and *must*, but with the present tense meaning (Hopper & Traugott [1993] 2003: 56).

Below, I focus on the rise of a new category of verbs, that of auxiliary ones, in ME from the perspective of reanalysis (cf. Lightfoot 1979; 1991, Warner 1993, Hopper & Traugott [1993] 2003, although these sources offer different views on the nature of the development of auxiliary verbs, including modal (auxiliary) verbs at the time). The rise of the category of auxiliary verbs, including modal (auxiliary) verbs, constitutes a case of recategorialization, i.e. a change in the category or status from lexical verbs to auxiliary verbs, in addition to changes in constituency and hierarchy status (Hopper & Traugott [1993] 2003: 56, 58). Hopper & Traugott ([1993] 2003: 55-58) argue that between the late 14th and the 15th centuries, the following types of constructions were available for verbs that were to become auxiliary verbs: a) question inversion and negation were without *do*; and b) the predecessors of ModE modal (auxiliary) verbs functioned as independent verbs, as in (29) below, and with the *to*-infinitive, as in (30) below, but also occurred in non-finite forms, as in (30) below, and in sequences, as in (31) below.

- (29) *She koude muchel of wandryng by the weye*
 She knew much of wandering by the way
 (c1390 Chaucer *Canterbury Tales Prologue* A.467; Hopper & Traugott [1993] 2003: 56; translation in the original)

‘She knew a lot about travel’

- (30) *any man þe whiche hadde mowzt to*
 any man the which had [been] able to
scapen þe deth
 escape the death
 (a1380 *Wycliffite Bible* 2 Par. 20.24; *MED* s.v. *mouen* v. 11b, as cited in Hopper & Traugott [1993] 2003: 56; translation in the original)

‘any man who had been able to escape death’

- (31) *No-þing to hafe is sum-tyme of need,*
No-thing to have is sometime of need,
bot nozt to may will haue is of grete
but not to able to will have is of great
vertew
virtue

(1434 Richard Misyn *The Mending of Life* 128/8; *MED* s.v. *mouen* v. 10a, as cited in Hopper & Traugott [1993] 2003: 56; translation in the original)

‘To have nothing is sometimes a necessity, but to desire [lit. to be able to will to have] nothing is a great virtue’

Therefore, the predecessors of auxiliary verbs were characterised by the occurrence in both finite and non-finite contexts, with direct object NP or complement clause, and in combination with another verb that was to become a modal (auxiliary) verb. However, by the 16th century, a set of individual changes (morphosyntactic and semantic changes outlined in the preceding paragraph) had led to a gradual disappearance of the above transitional constructions, and, together with other changes going on elsewhere, such as word order changes, has advanced the reanalysis of what was initially one category of verbs as two: lexical verbs and modal (auxiliary) verbs (Hopper & Traugott [1993] 2003: 57-58). Lightfoot (1979: 122) sees the development of modal (auxiliary) verbs as consisting of two stages,

[t]he first part of the story tells of complexity being developed gradually with an increasing number of exception markers. The second part of the story is a sudden, cataclysmic, wholesale restructuring of the grammar [i.e. reanalysis] whereby the exceptionality is, in a sense, institutionalized and the derivational complexity is eliminated at a stroke.

Although the available evidence points to recategorialization (i.e. reanalysis) having taken place at a deep structure level, the outcome of which was the emergence of the new category of verbs, that of auxiliary ones, there are a number of problems with the above story. Firstly, the individual disappearance of non-finite constructions related to the emergence of auxiliary verbs, in particular modal (auxiliary) verbs, is difficult to date. Fischer & van der Wurff (2006: 148; emphasis in the original), for example, argue that “[t]he evidence would have been more convincing if, due to the category change to Aux, the modals began to occur with *new* constructions, but this

is not the case.” Furthermore, there is also a question of whether the loss of the kinds of constructions outlined in the preceding paragraphs is unrelated and accidental. Warner (1993: 152) and Hopper & Traugott ([1993] 2003: 56), for example, argue that the group of verbs that became modal (auxiliary) verbs had already shown a number of morphosyntactic and semantic characteristics in OE that separated them from the other verbs. Warner (1993: 195) claims that “[t]he fact that [the] morphology [of the preterite-present verbs] was nonprototypical for verbs must also have encouraged (or been a precondition for) reanalysis”, whereas Fischer & van der Wurff (2006: 149) maintain that, “in OE the past-tense modals could be used to express present-time modality, so in that respect they differed from ‘normal’ verbs”. Finally, Lightfoot’s (1979) proposal suggests that the change from lexical to functional in relation to modal (auxiliary) verbs, and so auxiliary verbs, too, was radical. However, this view has led to reactions from others, arguing that the change was gradual (Plank 1984, Lightfoot 1991, Warner 1993, Hopper & Traugott [1993] 2003). Hopper & Traugott ([1993] 2003: 57), for example, demonstrate that a number of verbs in ModE are still undergoing some of the changes, with some of *be to*, *dare to*, *need to* and *ought to* either requiring or not requiring *do* support in negatives and questions.

Overall, therefore, a new category of verbs arose through reanalysis during the transition from OE to ME, namely that of auxiliary verbs. However, verbs belonging to this group are characterised by different degrees of auxiliiation. A number of features typical of this group of verbs can also be seen in the behaviour of OE *onginnan* and OE and eME *aginnan*, suggesting that both verbs have undergone reanalysis, too, as part of the process of auxiliiation. This will be discussed in the following section.

4.4.2.2 *Reanalysis of Onginnan in Old English, and Gan and Can in Middle English and Scots*

The preference of *gan* and *can* for the plain infinitive in ME and Scots, including the NME and ESc data sets, suggests that this verb and its variant follow the pattern of auxiliary verbs in terms of reanalysis; on the other hand, the occurrence of *beginnen*

with the (*for*) *to*-infinitive and the *at*-infinitive, in addition to zero complementation and other types of non-infinitival complements (both finite and non-finite forms) in these varieties and data sets suggests a split and reanalysis, but also gradualness of the change. A number of sources maintain that by selecting the plain infinitive, rather than the *to*-infinitive, *gan* and *can* have undergone a change from lexical to functional and so, this being one of the characteristics explaining the auxiliary status of this verb and its variant (I deal with the semantics of *gan* and *can* in Chapter Five). Evidence of reanalysis can also be found in relation to the predecessors of this verb and its variant in OE. For example, in relation to OE *onginnan* and OE *beginnan*, Los (2000; 2005: 82-99) proposes that in Ælfric, the choice between the plain infinitive and the *to*-infinitive found with these verbs depends on syntactic factors, but, at the same time, affecting the semantics of the verbs with which these types of complementation occur. She argues that both OE *onginnan* and OE *beginnan* with infinitive have different uses that are the result of the interplay of semantic and aspectual factors between the two infinitives, as well as discourse effects found in narrative contexts. She proposes that when the semantically bleached variant of either of the above-mentioned verbs is used, this determines the use of the plain infinitive, which in turn points to such constructions performing a grammatical function and so being grammaticalized. In this respect, Los (2000: 251) recognises the importance of the mechanism of reanalysis, but also highlights the significance of semantic bleaching, which follows from reanalysis (cf. Los 2005: 88-99),

[t]here are syntactic and semantic differences between the two constructions after these inchoative verbs which seem to point to auxiliary status when these verbs are followed by the bare infinitive, and lexical status when they are followed by the *to*-infinitive. As an auxiliary, *onginnan/beginnan* with the bare infinitive is shown to have been subject to considerable semantic bleaching. When *onginnan/beginnan* are followed by a *to*-infinitive, there is no bleaching: they are full lexical verbs with inchoative meaning.

Los (2000: 271) concludes that both OE *onginnan* and OE *beginnan* show auxiliary behaviour when occurring with the plain infinitive: a) “they can either follow or precede the matrix verb”; and b) “they are transparent to the argument structure of e.g. impersonal verbs, as if they had no argument structure of their own” (cf. Los 2005: 88-99). She also argues that such structures are seldom found in a V1 construction when governing the plain infinitive, this being due to the fact that the

events denoted by the plain infinitive are not segmentable. On the other hand, she claims that when both OE *onginnan* and OE *beginnan* are found with the *to*-infinitive, they refer to a segmentable and interruptible event. However, while I do not think that it is possible for these verbs to be both semantically empty and semantically full at the same time when occurring with either the plain infinitive or the *to*-infinitive, Los (2000; 2005) provides evidence for a crucial characteristic of grammaticalization, namely gradualness – as a result of this feature, the individual members of both groups can differ greatly in relation to the degree of grammaticalization. Furthermore, Los (2000: 271; 2005) finds that the plain infinitive found with OE *onginnan* and OE *beginnan* is not preferred when these verbs occur clause-initially. She demonstrates that this is due to the fact that not just the *þa* V pattern, but also V1 one, can function as an episode boundary marker, the difference between them being that the latter is responsible for signalling discontinuity, namely changes in events or turning points. Overall, Los (2000; 2005) concludes that the choice of infinitive with OE *onginnan* and OE *beginnan* is carefully selected in Ælfric – the more semantically bleached these verbs are, the more likelihood that they will occur with the plain infinitive. In this respect, it is possible to speak about the above items being grammaticalized constructions when occurring with the plain infinitive.

I agree with Los (2000; 2005) that the presence of the plain infinitive may be indicative of the strengthening of the degree of grammaticalization of OE *onginnan* and OE *beginnan* (cf. Brinton 1988: 117 on the development of the above from Germanic). However, there are also a number of other indications from her data that when found with the plain infinitive, as opposed to the *to*-infinitive, the above verbs are more advanced in the divergence of their development towards auxiliation. I would argue that there is a split with reference to hierarchical structure and category labels, resulting in bleached OE *onginnan* and OE *beginnan* showing a greater tendency for auxiliation when found with the plain infinitive. This kind of situation is described as layering, i.e. “[t]he persistence of older forms and meanings alongside newer forms and meanings, whether derived by divergence from the same source or

by renewal from different sources” (Hopper & Traugott [1993] 2003: 124).⁷² In other words, while new layers occur within a functional domain, the older ones remain alongside the new ones, interacting with them (Hopper 1991: 22), rather than competing with them or filling in potential gaps (Hopper & Traugott [1993]: 126). Finally, Hopper & Traugott ([1993] 2003: 126) argue that another characteristic of layering is the presence of contexts where the constructions in question involve a pragmatic difference, either clear or not. With reference to OE *onginnan* and OE *beginnan* in Ælfric, we can, therefore, find layering in that depending on the infinitive, instances of these verbs are either bleached (auxiliary uses), or not (ingressive uses).

Fischer, van Kemenade, Koopman & van der Wurff (2000: 218) argue that “it is only through the emergence of what would earlier have been ungrammatical constructions that we can obtain evidence for reanalysis.” Every instance of reanalysis, therefore, involves abduction (Hopper & Traugott [1993] 2003: 52) or is closely related to it (Fischer, van Kemenade, Koopman & van der Wurff 2000: 218, cf. Andersen 1973, McMahon 1994). Hopper & Traugott ([1993] 2003: 52) describe abduction, “[i]n some contexts two interpretations were possible, that is, there was at least the potential for ambiguity (also called ‘opacity’) that allowed for the structure to continue to be analysed as before, and for a new analysis to be innovated, and then to coexist with the earlier analysis” (cf. Diewald 2002: 116, Diewald & Smirnova 2012: 113-114, 116-117, 125). Fischer, van Kemenade, Koopman & van der Wurff (2000: 218) demonstrate a case of reanalysis, whereby the benefactive dative, originally governed by the main verb, as in (32) below, has been reinterpreted to function as the subject of the infinitival clause, as in (33) below. They also argue that this reanalysis is characterised by abduction in that a child guesses at a set of rules that have resulted in a particular construction. Crucially, “[t]he child may deduce the rule that

⁷² Hopper & Traugott ([1993] 2003: 126) provide other characteristics of layering: a) “grammaticalization [usually] does not result in the filling of any obvious functional gap. On the contrary, the forms that have been grammaticalized compete with existing constructions so similar in function that ... there is no obvious gap to be filled”; b) “[d]uring any phase of coexistence there are some contexts in which the two (or more) types in question involve a clear pragmatic difference. There are other contexts in which the choice between them is less clear with respect to pragmatic difference”; and c) “[q]uite often the newer layers of functionally similar constructions are symptomatic of more global adjustments.”

corresponds to the older structure, but she may also postulate an innovative structure ... which produces the same surface structure but also some ‘wrong’ or certainly innovative results” (Fischer, van Kemenade, Koopman & van der Wurff 2000: 218). With reference to Los (2000; 2005), I would maintain that during the process of grammaticalization of OE *onginnan* and OE *beginnan*, these verbs would probably have been characterised by abduction, too, resulting in ambiguity or opacity.

- (32) NP V NP [CP[IP PRO to V]]
 (‘benefactive construction’, traditionally called ‘organic *for*’)
- (33) NP V [CP[IP NP to V]]
 (‘subject construction’, or ‘inorganic *for*’)

That also the occurrence of *gan* and *can* with the plain infinitive in ME and Scots (in addition to semantic bleaching) may be seen as indicative of a greater degree of grammaticalization towards auxiliation, in view of claims in the literature that the more grammaticalized variant takes the plain infinitive, is discussed by Fischer (1992). Referring to verbs performing an auxiliary role, Fischer (1992: 318) argues, “it seems clear that the more grammaticalised the matrix verb is, or, in other words, the emptier it is of referential meaning, the more likely it is that the bare infinitive is found”. She adds that amongst such verbs are *gan* and *beginnen* in ME, too, maintaining, “*bi(ginnen)*, once it has developed into a kind of aspectual auxiliary or a purely periphrastic verb (especially in the past tense), normally appears with a bare infinitive” (Fischer 1992: 318), thus proposing an auxiliary status for these verbs, given that, “we do not expect a bare infinitive to appear where such grammaticalization is out of the question”.

However, not all sources argue for the auxiliary status of *gan* on the grounds of its occurrence with the plain infinitive. For example, Warner (1993: 139) draws attention to the fact that this verb does not exclusively occur with one type of infinitive or the other, despite being seen as a grammaticalized construction, “even uses of *ginnen* which have been classified as ‘auxiliary’ on semantic grounds seem to lack a restriction to the plain infinitive ... [in] meaningless ‘auxiliary’ uses ... even here there is variation between the *to*-infinitive and the plain infinitive” (cf. *MED* s.v. *ginnen* v. 3b). Nevertheless, such occurrences of *gan* with the *to*-infinitive could be

seen as instances of renewal, which is described by Hopper & Traugott ([1993] 2003: 122) as follows,

[i]n divergence existing forms take on new meanings in certain contexts, while retaining old meanings in other contexts [. However,] existing meanings may take on new forms [, which is referred to as] renewal. Renewal results primarily in alternate ways of saying approximately the same thing, or alternate ways of organizing linguistic material. Often, but not always, these new ways are periphrastic, i.e., phrasal.

The occurrence of *gan* with the *to*-infinitive as a case of renewal would presumably be due to the fact that the *to*-infinitival complement, already a possible complement after OE *onginnan* (Los 2000, 2005: 88-99), i.e. the predecessor of ME *gan* and *can*, came to be gradually used as a way to renew the ingressive meaning after OE *onginnan* has undergone semantic bleaching (cf. Los 2000, 2005: 88-99 on the grammatical status of etymologically related OE *beginnan*, especially when found with the plain infinitive). I will come back to the issue of renewal when I discuss the only instance of *gan* with the *to*-infinitive found in the NME data set.

While Los' (2000; 2005) investigations mainly discuss the auxiliation of OE *onginnan* and OE *beginnan* being dependent on the distinction between the two types of infinitive and the extent of bleaching (and whether the above verbs are in V1 or V2 word order), Sims' (2008) study focusses on the degree of grammaticalization of both OE *onginnan* and ME *gan*, therefore highlighting the etymological relationship between both verbs. For example, like Los (2000; 2005), Sims (2008: 122) seems to argue that OE *onginnan* already displays characteristics of reanalysis, "[*onginnan*] is still analyzed as a lexical element but is beginning to be used as a semantically weak form in periphrastic constructions. As rightward movement along the grammaticalization cline continues, non-aspectual *-ginnan* [i.e. *gan*] will be reanalyzed as a functional element."⁷³ However, unlike Los (2000; 2005), Sims (2008: 170) maintains that the process of grammaticalization is not complete for this verb in OE; she demonstrates that *onginnan* shares a number of syntactic characteristics with the group of verbs that were to become auxiliary verbs. For example, she argues that as opposed to *beginnan*, there are a number of instances of

⁷³ Cf. Halliday (1961), Heine (1992), Lehmann (1995) and Hopper & Traugott ([1993] 2003) for different terminology on the notion of the grammaticalization cline.

onginnan with the plain infinitive “in impersonal constructions with non-nominative subjects” (Sims 2008: 170, cf. Warner 1993: 103). She also claims, “*onginnan* occurs either before (Vf/Vnf) or after (Vnf/Vf) in verbal complements [and that it] occurs with *ge*-prefixed complements” (Sims 2008: 170), which, in a similar way to the predecessors of ModE modal (auxiliary) verbs, would indicate, “*onginnan*’s syntactic behaviour is more auxiliary-like” (Sims 2008: 130, cf. van Gelderen 2004). This leads Sims (2008: 170) to conclude that “these differences in syntactic behaviour align *onginnan* with the early pre-auxiliary class ... [h]owever, *onginnan* retains its semantic content throughout the OE period”.

Sims (2008: 171) also reaches some important conclusions about the verb *gan* in ME. She argues that during this period, the phonologically reduced *gan* is on the increase. Because it is also consistently used by *Lazamon* in a Vf/Vnf pattern, whereas *agon*, the preterite form of OE and eME *aginnan*, in a Vnf/Vf pattern, this leads her to speculate that “his use of [the Vf/Vnf] pattern [with *gan*] indicates that he considered it grammatically acceptable” (Sims 2008: 171). Finally, given that *gan* reaches its highest point between 1300 and 1350 in terms of frequency of occurrence (Terasawa 1974), Sims (2008: 171) concludes, “[f]rom this time on, *gan* occurs mainly with the bare infinitive, which, again, patterns more closely with the pre-auxiliary (modal) class. We see then that as the lexical use of *gan* decreases, the more auxiliary-like use of *gan* increases.” Overall, therefore, Sims (2008) argues that although OE *onginnan* shares a number of syntactic properties with the group of verbs that were to become auxiliary verbs, this verb is characterised by “instability”, i.e. change, as opposed to the “stability” of *beginnan*, i.e. being “at a historical dead end”; this ‘instability’ of OE *onginnan* is mainly due to the fact that instances of this verb only begin to “be used as a semantically weak form” (Sims 2008: 122). She then maintains that structurally, the reanalysis of the verb *gan* with the plain infinitive takes place at the beginning of the ME period, patterning this verb more closely to the “pre-auxiliary (modal) class” (Sims 2008: 171). Although Sims’ (2008) proposals about OE *onginnan* and ME *gan* are in line with the theory of grammaticalization, whereby both verbs become grammatical constructions by undergoing reanalysis (and, eventually, in relation to *gan*, complete semantic

bleaching), her claims differ from those by Los (2000; 2005), who essentially argues both reanalysis and semantic bleaching for instances of OE *onginnan* and OE *beginnan* with the plain infinitive.

There is also another interesting point to note about Sims' (2008) proposals with regards to the grammaticalization of OE *onginnan* and ME *gan*. Sims (2008: 171) argues that there are two separate types of reanalysis occurring with these verbs: the first one relates to the structure of OE *onginnan* with the plain infinitive, given that this verb is beginning to be used as a periphrastic construction in OE, alongside instances of this verb as a lexical verb; the second one, on the other hand, concerns the aphetic form *gan* with the plain infinitive as a fully-fledged functional element in ME. However, although a suggestion that *gan* undergoes further reanalysis would strengthen the claim that this construction is becoming more grammaticalized, i.e. becoming a member of the "pre-auxiliary (modal) class" in ME (Sims 2008: 171), there might be a concern that the type of evidence that Sims (2008) uses to support her claims to determine the degree of grammaticalization of this verb would influence the outcome. For example, Sims (2008) notes that prosodic features of a text may influence the choice of *gan* with either the plain infinitive or the *to*-infinitive, this claim being similar to metrical proposals outlined in Section 4.4.1. I would also argue that the Vf/Vnf pattern Sims (2008) identifies that *Lazamon* might prefer to use with *gan* is indicative of this verb being adjacent to its complement in a line of verse, rather than showing a degree of bonding, i.e. new constraints on the construction, where, for example, a linguistic element such as an adverb may not intervene between the main verb and the verb in infinitive. Furthermore, the Vf/Vnf pattern might also reflect general syntactic properties of ME at the time, especially if other verbs that take infinitival complements have the infinitive adjacent at least as often as the verb *gan*.

While *gan* and *can* with the plain infinitive differ from lexical verbs in that they undergo reanalysis and semantic bleaching, some sources suggest that instances of the predecessors of this verb and its variant show bleached uses in OE when found

with the *to*-infinitive, too.⁷⁴ For example, Brinton (1988: 159-161) and Sims (2008: 121-122) provide a single instance of what they claim is a semantically bleached example of OE *onginnan* with the *to*-infinitive from Ælfric, as in (34) below, beside a number of instances of this verb with the plain infinitive, as in (35) below. Brinton (1988) copies this example of OE *onginnan* with the *to*-infinitive in (34) below from Calloway (1913: 53), who shortens it to illustrate inflected forms of the *to*-infinitive found with this verb, rather than being interested in the semantics of such constructions (this example is provided in full in (36) below). Sims (2008: 121), in turn, duplicates this example from Brinton (1988: 160), arguing that “[t]he non-aspectual [i.e. non-ingressive] use of *onginnan* ... is extremely minimal in OE ... with only a few individual cases appearing in the manuscripts”. On the basis of this example of OE *onginnan* with the *to*-infinitive, as in (34) below, and with the plain infinitive, as in (35) below, she maintains that, “[b]ecause non-aspectual *onginnan* has few, if any, semantic features, the form is not relevant to the argument structure, and is, therefore, more auxiliary-like in its function” (Sims 2008: 122). Overall, therefore, both Brinton (1988) and Sims (2008) suggest that instances of OE *onginnan* reanalyse with the plain infinitive and the *to*-infinitive, and undergo semantic bleaching to be more auxiliary-like in this verb’s function.

- (34) *Witodlice* ... *ongann* *se* *hiredes* *ealdor to*
 Truly ... did the household’s elder to
agyldenne *þone pening*
 repay the penny
 (ÆCHom II, 5 46,137; Brinton 1988: 160)

‘Truly, the lord of household repaid the wages’⁷⁵

- (35) *ge* *þa* *sciran miht* *deman ongunnon*
 you then bright might judge did

⁷⁴ It has been argued that *beginnen* can also emphasise the occurrence of an event or action denoted by the verb in the plain infinitive or the *to*-infinitive (MED s.v. *biginnen* v. 6, Häusermann 1930: 21-22, Koziol 1932: 133-134, Visser 1969: 1580-1581, Sims 2008: 185-186). However, all of Visser’s (1969: 1580) examples of *beginnen*, probably having a function of the periphrastic auxiliary, are with the (*for*) *to*-infinitive, in addition to the difficulty of identifying whether in particular cases this verb has stopped to refer to “the initial stage, the commencement, of the action.” Visser (1969: 1580) concludes, “[t]he subjoined evidence is, admittedly, of a very tentative character, contextual clues for determining the ‘correct’ interpretation being either missing or inconclusive.”

⁷⁵ Brinton (1988: 160) translates the above as ‘Certainly repaid (NOT began to repay) the elder of the house the penny’; I have provided an alternative translation.

(*El* 310b-311a; Brinton 1988: 160; translation in the original)

‘you then judged (NOT began to judge) the bright power (of Christ)’

However, Los (2000: 259) convincingly argues against the claim that OE *onginnan* showed bleached uses when used with the *to*-infinitive in Ælfric’s works. She claims that Brinton’s (1988) and Sims’ (2008) example in (34) above, quoted in its entirety in (36) below, has, in fact, *onginnan* used in an ingressive sense. This interpretation can be reached on the basis of the context of a discussion of the Parable of the Workers of the Vineyard. In this parable, Jesus explains that anyone who accepts the invitation to work in the vineyard, representing the Kingdom of Heaven, will obtain a day’s wages, i.e. the reward of one penny (chiefly used in Biblical use and translations, the NP *pening* is a rendition of Classical Latin *dēnarius*, referring to a Roman coin of low denomination, or, especially in earlier use, a denarius; *OED* s.v. *penny* v. 2a). Los (2000: 259) maintains that the action or event presented here is iterative, with the owner of the vineyard paying out a multiplicity of pennies; although the NP *pening* refers here to a singular penny, it denotes an unspecified amount in this example. In this respect, the situation presented in (36) below is durative, i.e. non-punctual (the effects of the Aktionsart of the verb in infinitive on the meaning of OE *onginnan* and OE and eME *aginnan*, as well as ME *gan* and *can* will be addressed in Chapter Five). Los (2000: 260) concludes that only instances of OE *onginnan*, with the plain infinitive undergo reanalysis and semantic bleaching, displaying auxiliary uses in Ælfric,

[t]here are ... no examples of such ‘pleonasm’ with a *to*-infinitive. A final point to note here is that Ælfric himself, in his Latin Grammar, systematically employs *onginnan* + *to*-infinitive to translate Latin inchoatives, and not the bare infinitive, e.g. *calesco*, *ic onginne to wearmigenne* ‘I begin to become warm’ (<ÆGram 212.3> ...). This further supports the idea that only the *to*-infinitive expresses strong ingression.

(36) *Witodlice* *fram* *ðam* *endenextan* *ongann* *se*
Truly from the last-ones began the
hiredes *ealdor to* *agylденne* *þone pening*
household’s elder to pay the penny
(*ÆCHom* II, 5 46,137; Los 2000: 259)

‘Truly, the lord of the household began to pay the wages starting from the last ones’,⁷⁶

Therefore, the reason why Los (2000) can make claims that OE *onginnan* and OE *beginnan* with the plain infinitive can no longer express the idea of ingression by Ælfric’s time is the incompatibility of these verbs and their complements with the V1 constructions. The V1 construction in the works by Ælfric is equivalent to ModE ‘just when’, requiring at the same time an interruptible action. As it is possible to find instances of OE *onginnan* with the *to*-infinitive in such cases suggests that such temporal segmentation can presumably no longer be provided by OE *onginnan* and OE *beginnan* with the plain infinitive. This interplay of syntactic, semantic and textual factors leads to the process of grammaticalization of OE *onginnan* with the plain infinitive. It also has consequences on the development and status of the verb *gan* and its variant *can* in the NME and ESc data sets, as discussed below.

4.4.2.3 *Development of Gan and Can, as well as Beginnen in Northern Middle English and Early Scots*

4.4.2.3.1 *Gan and Can*

An important point that can be taken from the literature review in the preceding section is that OE *onginnan*, the predecessor of *gan* and *can* in ME and Scots, reanalyses with the plain infinitive and, as a result of semantic bleaching, already has an auxiliary function by the time of Ælfric (Los 2000; 2005). The findings in Section 4.3.1 demonstrate that the verb *gan* shows a preference for the plain infinitive in the NME data set, as in (37) below, with one instance of this verb also being recorded with the *to*-infinitive there. Its variant *can*, on the other hand, is exclusively used with the plain infinitive in the NME and ESc data sets, as in (38) below. Therefore, relying on the view that the *to*-infinitive was gradually used with verbs during the ME period is not sufficient in relation to *gan* and *can* here.

(37) *To the schyp sche come full tye,*
 To the ship she came full quickly,

⁷⁶ Los (2000: 259) translates the above as ‘Truly, from the last ones began the lord of the household to pay the penny’; I have provided an alternative translation.

And on the syde gan he smyte:
 And on the side did he strike:
The lady gan up stonde.
 The lady did up stand.
 (1420-1500 *Sir Eglamour of Artois* l. 877-879; TEAMS)

‘She came to the ship very quickly, and he struck on the side: the lady [then] stood up.’

(38) *Florent to the stede can gone;*
 Florent to the place did go;
So feyre an hors sye he never none
 So fair a horse saw he never none
 (1420-1500 *Octavian* l. 724-725; TEAMS)

‘Florent went to the place; never before did he see such a beautiful horse’

Below I outline the development of *gan* and *can* in the NME and ESc data sets from the point of view of reanalysis and analogy, highlighting the effects of these mechanisms on the structure of this verb and its variant – the changes concerning both items under investigation are schematised in Figure 1. However, it will be recalled from Section 1.3.2 that although there is some textual evidence available in the Northumbrian dialect of OE (Cuesta & Silva 2008: 50-51), this is limited; furthermore, there is no literary evidence available in Pre to literary Scots, making it difficult to establish developments with regards to the predecessors of *gan* and *can* in the north (Macafee 1992/93: §1.4, King 1997: 156). For these reasons, I use here textual evidence from southern dialects of OE and eME, bearing in mind the similarity, but also the differences, between southern and northern dialects. Another matter that should be mentioned here again concerns the representativeness of the data sets used in this study, especially that the occurrence of this verb and its variant predominantly or exclusively with the plain infinitive, but not with the *to*-infinitive, could be due to accidental gap in the data sets used here.

Figure 1

Syntactic Development of *gan* and *can* with Infinitive in the NME and ESc Data Sets

Stage I

By reanalysis⁷⁷

[*onginnan/aginnan*] ->

[*onginnan/aginnan* [plain infinitive]]
(non-punctual verb)
(*onginnan/aginnan* continue to operate as independent verbs)

Stage II

By analogy

[*onginnan/aginnan* [plain infinitive]] ->

(non-punctual verb)

[*onginnan/aginnan* [plain infinitive]]
(punctual (achievement) verb)

By analogy

[*onginnan/aginnan* [plain infinitive]] ->

(non-punctual verb)

[*onginnan/aginnan* [*to*-infinitive]]
(non-punctual verb)

Stage III

By ?further reanalysis

[*onginnan/aginnan* [plain infinitive]] ->
(punctual (achievement) verb)

[*gan/can* [plain infinitive]]
(punctual (achievement) verb)

Stage IV

By analogy

[*gan* [plain infinitive]] ->
(punctual (achievement) verb/non-punctual verb)

[*gan* [*to*-infinitive]]
(punctual (achievement) verb/non-punctual verb)

⁷⁷ Traugott & Trousdale (2013: 36) convincingly argue that there is a terminological problem with 'reanalysis', "[i]f a language user who has not yet internalized the construction in question, interprets a construction in a different way from the speaker, 're'-analysis has not occurred, only 'different' analysis; strictly speaking, one cannot 're'-analyze a structure one does not 'have'; instead, they propose the term 'neoanalysis'. However, although I agree with this claim, I continue to use the term 'reanalysis' for the sake of clarity.

Stage I in the development of *gan* and *can* with infinitive in the NME and ESc data sets initially involves the reanalysis of one category of OE *onginnan* and OE and eME *aginnan* as two: as an independent verb, as in (39) below, and with infinitive, as in (40) below and (41) below. The above can be supported with findings from the literature, where, for example, Ogura (1997: 406-407) provides evidence of occurrence of *onginnan* and *aginnan* as an independent verb and with infinitive in various OE and eME texts. To summarise her findings: a) the verb *onginnan* in the preterite is used in 856 instances with infinitive and in 77 instances as an independent verb; b) this verb in what Ogura (1997: 406) refers to as “Other Forms” occurs in 82 instances with infinitive and in 119 instances as an independent verb; c) the verb *aginnan* (including the form *anginnan*) in the preterite occurs in 56 instances with infinitive and in 19 instances as an independent verb; and d) this verb is found in “Other Forms” in seven instances with infinitive and in 17 instances as an independent verb.⁷⁸ Overall, therefore, even though the occurrence of both verbs in the preterite may be due to the nature of the textual material, especially that poets often used both the present and the preterite to signal various stages in the narrative (Richardson 1991a), and that the above represent ‘raw’, rather than normalised, frequencies, the above findings by Ogura (1997: 406) show that OE *onginnan* and OE and eME *aginnan* reanalyse as two separate categories: an independent verb and with infinitive. Furthermore, in the preterite, these verbs predominantly occur with infinitive, rather than as independent verbs, although in other forms, functioning as independent verbs prevails over these verbs found with infinitive.

(39) *Incipit liber ... onginþ seo boc peri didaxeon*
 Incipit liber ... begins the book Peri Didaxeon
 (c1150 *Peri Didaxeon* 3/3; *MED* s.v. *aginnen* v. 3)

‘The book *Peri Didaxeon* begins’

(40) *Ongan ceallian þa ofer cold wæter*
 Began call then over cold water
 (Mald 94; Brinton 1988: 159; translation in the original)

‘[He] began to call out then over the cold water’

⁷⁸ Ogura (1997: 408) also provides frequencies of occurrence of OE and eME *ginnan*, showing that it is found with infinitive in 572 instances (in *Lambeth Homilies*, *Ormulum* and *Lazamon’s* works), with two instances of <geginnen> with (the *to-*) infinitive from *Lazamon* excluded in her counts.

- (41) *Agan se cyng georne to smeagenne wið his*
 Began the king eagerlyto discuss with his
witan
 advisers
 (ChronE [Plummer] 1006.37; Brinton 1988: 118)

‘The king eagerly began to consult with his council’⁷⁹

In relation to Stage II, it is possible to identify analogical change at work, whereby the occurrence of verbs in infinitive with OE *onginnan* and OE and eME *aginnan* involves punctual (achievement) verbs, as in (42) below and (43) below, which are usually incompatible with the ingressive meaning, originally associated with the above verbs (the meaning of *gan* and *can* will be addressed in Chapter Five). Since there is bleaching involved and given that it is the result of reanalysis, or as Harris & L. Campbell (1995: 92) rather put it, “it is the essence of the reanalysis itself”, it appears that there is also a gain in the functional domain for OE *onginnan* and OE and eME *aginnan*.

- (42) *ge þa sciran miht deman ongunnon*
 you then bright might judge did
 (El 310b-311a; Brinton 1988: 160; translation in the original)

‘you then judged (NOT began to judge) the bright power (of Christ)’

- (43) *Þa þe cniht wende þat hit þe eotend*
 Then the knight weened that it the giant
weoreand he an-bursten agon swulc weore a wilde
 were and he enrage did so were a wild
bar
 boar
 (Lazamon *Brut* 25831-25834; Brinton 1988: 158)

‘Then the knight thought that it was the giant and he enraged like a wild boar’⁸⁰

⁷⁹ Brinton (1988: 118) translates the above as ‘The king began eagerly to rule with his advisors’; I have provided an alternative translation.

⁸⁰ Brinton (1988: 158) translates the above as ‘then the knight went to the place where the giants were and he burst out (NOT began to burst out) just as if he were a wild bear’. However, the noun *bar* refers to ‘an uncastrated male swine (either wild or domesticated)’ (*MED* s.v. *bōr* n. 1), with the expression *brust (iburst) as a bor* meaning ‘bristly as a boar, bristling (or showing anger) like a boar’ (*MED* s.v. *bōr* n. 1d), rather than a bear. I have provided an alternative translation.

Such bleached instances of OE *onginnan* and OE and eME *aginnan*, as in (42) above and (43) above, respectively, only occur with the plain infinitive (cf. Los 2000: 259 on OE *onginnan* and OE *beginnan* with the plain infinitive in Ælfric). In view of claims in the literature that the more grammaticalized variant takes the plain infinitive, this is another indication of the increasing grammaticalization of OE *onginnan* and OE and eME *aginnan* –Warner (1993: 136, 137) maintains that in OE, both the plain infinitive and the *to*-infinitive are found as complements with verbs, but the former was becoming a property of a small number of verbs, such as the predecessors of the ModE modal (auxiliary) verbs, even though it still had a much more wider distribution then, “avoidance of *to* in an infinitive complement was a remarkable restricted property in Old English, certainly by late Old English as the incidence of *to* increased”.

Looking at the findings by Ogura (1997: 406-407), it transpires that not only are OE *onginnan* and OE and eME *aginnan* more frequent with the plain infinitive, rather than the *to*-infinitive, but also that they are more commonly found with the plain infinitive when occurring in the preterite; the occurrence of both verbs in the preterite, therefore, appears to be significant for the emergence of *gan* and *can* in ME and Scots in the preterite in form: a) OE *onginnan* in the preterite is used in 814 instances with the plain infinitive and 42 instances with the *to*-infinitive; b) in other forms, it occurs in 78 instances with the plain infinitive and in four instances with the *to*-infinitive; c) OE and eME *aginnan* (including the form *anginnan*) in the preterite occurs in 53 instances with the plain infinitive and three instances with the *to*-infinitive; and d) in other forms, it is found in six instances with the plain infinitive and one instance with the *to*-infinitive. Despite these being ‘raw’ numbers, evident is a preference of OE *onginnan* and OE and eME *aginnan* for the plain infinitive, rather than the *to*-infinitive, especially when the meaning of these verbs is bleached.⁸¹

⁸¹ Similar findings are found in Callaway (1913: 279-287), but comprehensively and conveniently summarised in Sims (2008: 113; 2014: 63), who shows that 96% and 85% of all instances of OE *onginnan* and OE and eME *aginnan*, respectively, occur with the plain infinitive in OE texts, the remaining 4% and 15% of these verbs, respectively, being used with the *to*-infinitive. Sims (2008: 113; 2014: 63), therefore, shows that in Callaway’s (1913: 279-287) data, 977 and 37 instances of *onginnan* are used with the plain infinitive and the *to*-infinitive, respectively, while only 28 and 5

During Stage II in the development of *gan* and *can* with infinitive in the NME and ESc data sets, there is also extension by analogy of OE *onginnan* and OE and eME *aginnan* with the plain infinitive, as in (44) below, to these verbs with the *to*-infinitive⁸², as in (45) below. The extension by analogy of OE *onginnan* and OE and eME *aginnan* with the plain infinitive to these verbs with the *to*-infinitive is indicative that the newly grammaticalized *to*-infinitive continued to gain ground from the OE period onwards, except for a number of verbs that were still restricted to the plain infinitive (Warner 1993: 103, Fischer & van der Wurff 2006: 149). Furthermore, the occurrence of some instances of OE *onginnan* and OE and eME *aginnan* with the *to*-infinitive also suggests that this complement is used to reinforce the original ingressive meaning associated with both verbs before this ingressive meaning is bleached – Los (2000: 259-260) argues that in Ælfric, when OE *onginnan* and OE *beginnan* are used with the *to*-infinitive, both verbs have ingressive meaning.

- (44) *Ongan þa dryhtnes æ dæges ond nihtes ...*
 Did then God's law days and nights ...
georne cyðan
 eagerly exhibit
 (*El* 198a-9b; Brinton 1988: 161; translation in the original)

‘[He] then eagerly made known (NOT began to make known) the law of God day and night’⁸³

instances of *aginnan*, respectively. No instances of *aginnan* are recorded by Callaway (1913: 279-287) in Alfred and Ælfric, but in only what Sims (2008: 113; 2014: 63) refers to as “Other texts”. In a similar way, Los’ (2000: 256) findings show that a preference of OE *onginnan* for the plain infinitive is evident – 18 instances of this verb with the plain infinitive and seven instances of this verb with the *to*-infinitive.

⁸² As a separate type of change, there is the reanalysis of the *to*-infinitive. As already stated in Section 4.2, the *to*-infinitive originally derives from a preposition, although its function in OE is not that of a prepositional phrase anymore, but “a full-blown clause” (Los 2005: 191). Los (2005: 153-190) convincingly argues that during the OE period, the *to*-infinitive, which, unlike other nominal and prepositional objects, as well as finite clauses, occurred to the right of its governing verb, competing with the finite subjunctive clause, ousting it in different environments, “[a]ll the available evidence suggests, then, that the *to*-infinitive was already being analysed as a non-finite subjunctive in OE” (Los 2005: 189-190). Furthermore, the bond between *to* and its infinitive suggests that “*to* was a clitic or even a bound morpheme: nothing could intervene, and *to* could not be dropped from second (or subsequent) conjuncts in a sequence of coordinated *to*-infinitives” (Los 2005: 230).

⁸³ Wuth (1915: 57) provides the following German translation of this example: “fing an bei Tag und Nacht Gottes Gesetz zu verkündigen”, which can be transcribed as ‘began to preach the law of God day and night’.

- (45) *hine ongunnon ærst to torfienne*
 him began first to stone
 (ÆCHom I, 3 50.14; Brinton 1988: 160; translation in the original)

‘[they] began first to stone him’

During Stage III, when the use of OE *onginnan* and OE and eME *aginnan* with a punctual (achievement) verb in infinitive becomes conventionalised, there is, arguably, further reanalysis of the single morpheme *gan* and *can* with infinitive, such construction types being found in the NME and ESc data sets, as in (46) below and (47) below, respectively. This further reanalysis is following the weakening of the prefixes on OE *onginnan* and OE and eME *aginnan*, and the decategorialization of *gan* and *can* in the third person singular, which results in single morpheme forms in these data sets. It is also similar to the reanalysis of the auxiliary construction *be going to* as a single morpheme *gonna*, where it is rebracketed as [*gonna* + verb] (Hopper & Traugott [1993] 2003: 68). Because *gan* and *can* in auxiliary use no longer express ingressiveness and freely occur with punctual (achievement) verbs in infinitive, this further reanalysis also permits the occurrence of this verb and its variant with adverbials (or adverbs in Brinton 1983; 1988) of duration or iteration, as in (48) below. Overall, therefore, by this point, instead of focussing on the inception of an event or situation denoted by the verb in infinitive, such instances of the verb *gan* and its variant *can* in the NME and ESc data sets will portray this event or situation as a whole, i.e. complete, in reference to other situations or events in the context, despite a number of some instances of *gan* still being interpreted as ingressive, or also ambiguous, as is the case for this verb and its variant in the data sets.

- (46) *Pe woman þan hir dedes gan rew*
 The woman then her deeds did repent
And þus scho said vnto Jesu,
 And thus she said unto Jesus,
 (1350-1420 *The Northern Homily Cycle* P II p. 82; *HCET*)

‘The woman then repented her sins and said to Jesus in this way,’

- (47) *Wit þat stan he laid in sling,
 With that stone he laid in sling,
 Sua stalworthli he lete it suing
 So strongly he let it swing
 Þat in his frunt þat stan he fest,
 That in his forehead that stone he cast,
 Þat bath his eien vte can brest
 That both his eyes out did burst
 (1350-1420 *Cursor Mundi* p. 438; *HCET*)*

‘He put that stone in the sling, and let it swing so strongly that he cast it on the forehead, that both of his eyes came out’

- (48) *Ful wel war þai wit him þat night
 Full well were they with him that night
 Þe morn his ass þan can he dight,
 The morn his ass then did he saddle,
 To ierusalem he tok þe strete
 To Jerusalem he took the road
 (1350-1420 *Cursor Mundi* pp. 1011-1012; *HCET*)*

‘They were able to be with him that night. The next morning he saddled his donkey, and took the road to Jerusalem’

During Stage IV, however, there is also evidence of the verb **gan** occurring with the *to*-infinitive in ME, including one instance of this verb with this type of complementation in the NME data set, as in (49) below.

- (49) *The lyones fonde thay in hir dene,
 The lioness found they in her vale,
 A knave childe laye sowkand hir then
 A male infant lay sucking her then
 And gan with the lyones to playe.
 And began with the lioness to play.
 Umwhile the childe sowkede hir pappe,
 Sometimes the child sucked her breast,
 Umwhile gan thay kysse and clappe;
 Afterwards did they kiss and embrace;
 (1420-1500 *Octavian* l. 442-446; *TEAMS*)*

‘They [the men] found the lioness in her vale, a male infant lay sucking her then and began to play with her. Sometimes the child sucked her breast, afterwards, they kissed and embraced;’

The occurrence of the verb *gan* with the *to*-infinitive could be seen as a case of renewal, whereby “existing meanings may take on new forms” (Hopper & Traugott ([1993] 2003: 122), a way of obscuring the unidirectionality of the process of grammaticalization of *gan* with the plain infinitive. On the other hand, a point worth noting is that the verb *play* also existed in the prefixed form *geplægia* in the Northumbrian dialect of OE (*OED* s.v. *play* v.). Sims (2008: 75), for example, argues that the prefix *ge* was used in OE as a marker of perfective aspect, but van Gelderen (2004: 209, as cited in Sims 2008: 76) maintains that the loss of such prefixes results in a reanalysis of other forms, with the *to*-infinitive becoming associated with the notion of imperfective. Van Gelderen (2004: 2010, as cited in Sims 2008: 76) then argues that “[t]his reanalysis ... occurs at the end of the ME period and involves a shift from lexical aspect to grammatical aspect – the prefixes *ge-/for-* originate in ... lexical aspect ... but *-ing, to, have,* (and modals) are reanalysed as the head of [Aspect Phrase].” Nevertheless, the wider context and the fact that the verb in infinitive is an activity verb denoting duration (cf. Chapter Five) support the ingressive interpretation of *gan* with the verb *playe* in the *to*-infinitive in this example.

4.4.2.3.2 *Beginnen*

Reanalysis and analogy apply to *beginnen* with its complementation in the NME and ESc data sets, too, as schematised in Figure 2.

Figure 2

Development of *beginnen* with Infinitive in the NME and ESc Data Sets

Stage I

By reanalysis

[*beginnan*] -> [*beginnan* [plain infinitive]]
(non-punctual verb)
(*beginnan* as an independent verb
continues to operate)

Stage II

By analogy

[*beginnan* [plain infinitive]] -> [*beginnan* [*to*-infinitive]]
(non-punctual verb) (non-punctual verb)
(including other types of infinitive)
(also in the NME and ESc data sets)

Stage III

By analogy

[*beginnan* [*to*-infinitive]] -> [*beginnan* [*to*-infinitive]]
(non-punctual verb) (?punctual (achievement) verb)
(also in the NME data set)

However, unlike in the case of *gan* and *can*, there is no evidence of semantic bleaching in relation to *beginnen*, except for a small number of instances discussed below. Therefore, Stage I involves the reanalysis of one category of OE *beginnan* as two, with this verb occurring: a) as an independent verb, as in (50) below; and b) with the plain infinitive, as in (51) below.

(50) *Æfter ðám beginne se abbod*
After those begin the abbot
(*Benedictine Rule* 35, 22; Bosworth 2010 s.v. *be-ginnan* v. Add; I 2)

‘After those the abbot began’

(51) *he þæt westen ærest eardigean began*
he that desert first dwell began
(*Saint Guthlac* 4.26; *DOE* s.v. *be-ginnan* v. 1a)

‘he began to inhabit the desert at first’

Stage II is characterised by analogy, whereby OE *beginnan* begins to occur with the newly grammaticalized *to*-infinitive, and (52) below (including other types of infinitive in respective varieties), although instances of this verb with the plain infinitive are also still found. That the *to*-infinitive is gaining ground during the OE period, increasingly occurring with OE *beginnan*, too, is reflected in Callaway’s (1913: 279-287) findings summarised by Sims (2008: 113; 2014: 63): this type of infinitive occurs in 57 instances of *beginnan* in Ælfric and other texts (67%), as opposed to 28 instances with the plain infinitive there (33%); likewise, Los’ (2000: 256) findings show that the *to*-infinitive is used with this verb in 14 instances, whereas with the plain infinitive in seven (cf. Ogura 1997: 409). The occurrence of the plain infinitive and the (*for*) *to*-infinitive (including other types of infinitive in respective varieties) with this verb can also be observed in the NME data set in Period I, with significantly more instances of the latter than the former (93% vs. 4%; the remaining being the *at*-infinitive), with no instances of the plain infinitive in the NME data set in Period II and in the ESc data set in both periods.

(52) *ða* *began* *he* *to* *wyrcente* *wundra*
 then began he to work wonders
(Ælfric *Catholic Homilies: Feria III de Dominica Oratione: Clemoes*
I, 19 325.4; DOE s.v. *be-ginnan* v. 1a)

‘then he began to work wonders’

Finally, Stage III includes further analogy in the development of OE *beginnan* with infinitive, whereby, arguably, semantically bleached instances of this verb are found with punctual (achievement) verbs, resulting in the non-ingressive meaning of this construction in OE, as in (53) below (DOE s.v. *be-ginnan* v. 1c), but also *beginnen* in ME, as in (54) below (MED s.v. *beginnen* v. 6), and in the NME data set, as in (55) below.⁸⁴

(53) *þa* *began* *se* *preost swa* *swa* *he* *God* *lufode*
 then did the priest so as he God loved
 his *gebedu* *singan* and *swyðe* *fæstan*

⁸⁴ However, such instances of *beginnen* are classed as ‘ambiguous’ in this study, given that either a non-ingressive or an ingressive interpretation is possible in such cases.

his offices sing and very much
 (Ælfric *Lives of Saints: Passion of Saint Alban* 23; *DOE* s.v. *beginnan* v. 1c.i)

‘then the priest sang his offices very much because he loved God’

- (54) *Miht he noght blinne To bihold þat fair maydene,*
 Might he not stop to behold that fair girl,
Hou all hir dedis begane hir seme
 How all her deeds did her befit
 (a1400 *Cursor Mundi* 3311 (Göttingen MS Theol. 107); *MED* s.v. *biginnen* v. 6)

‘He cannot cease to behold that fair maiden, how all her deeds befitted her’

- (55) *and be þat at I sulde begyn to*
 and by that that I should do/begin to
lose þe vertue of my mekenes
 lose the virtue of my humility
 (1420-1500 *Alphabet of Tales* p. 509; *ICoMEP*)

‘and by that, that I would do/begin to lose the virtue of my humility’

With reference to ***begouth***, although it only occurs as an independent verb in the ESc data set, according to reference sources like the *DSL* (s.v. *begin* v. 1b b), this variant is also found with infinitive in Scots, as in (56) below.

- (56) *Thane begouth thai soroful to be*
 Then began they sorrowful to be
 (1380 *The Scottish Legendary* v.153; *DSL* s.v. *can* v. 1b b)

‘Then they began to be full of sorrow’

However, although formed on analogy to ***couth***, i.e. the preterite of ***can*** as a variant of ***gan***, ***begouth*** is found with the *to*-infinitive only in the *DSL*, as shown in the above example. This further highlights the difference in grammatical status between ***begouth*** and ***couth***, as well as the extent of grammaticalization of both verbs, although the lack of instances of this variant with the plain infinitive may be due to accidental gap in the ESc data set. Despite associations in form of ***begouth*** with ***couth***, the former is used with the type of infinitive (*to*-infinitive) different from that used by the latter (plain infinitive), as recorded in ESc (*DSL* s.v. *can* v. 2b, *OED* s.v.

can v.² II 4). Furthermore, there are also differences in meaning of *begouth* and *couth*, with the former retaining its ingressive meaning, whereas the latter being used with non-ingressive meaning (*DSL* s.v. *can* v. 2b). So, although nearly all instances of *couth* in the *DSL* are with activity verbs, there is one instance of this variant with a punctual (achievement) verb, as in (57) below, too. Whereas Taylor (1917: 579) lists such punctual (achievement) verbs as *clos*, *conquir* and *ban* amongst other types of verbs found with this variant.

(57) *The barde worth brane wod, and bitterly couth*
 The bard worth brain wood, and bitterly did
 ban
 curse
 (c1450-1452 Holland *The Book of the Howlat* 811; *DSL* s.v. *can* v.
 2b)

‘The bard became mad and cursed bitterly’

Overall, therefore, despite evidence of reanalysis and forms like *begouth*, the verb *beginnen* does not appear to show the same degree of change towards auxiliation in the NME and ESc data sets as *gan* and *can* do in these data sets. This is because not all reanalysis leads to grammaticalization (cf. Traugott & Heine 1991, Haspelmath 1999, L. Campbell 2001), one of the other indicators of grammaticalization taking place being semantic bleaching, as discussed in Chapter Five.

4.5 Summary

In this chapter, I presented an overview of changes in relation to the infinitive found with verbs in the transition from OE to ME. Although both the plain infinitive and the *to*-infinitive were used in ME, the application of the former was restricted to a small group of verbs, usually characterised by functional uses (modal verbs, verbs of physical perception, causatives), with other verbs like *pencan* ‘think’ in OE, and in ME also *helpen*, *maken*, as well as “the *ginnan*-verbs” taking both the plain infinitive and the *to*-infinitive (Fischer 1992: 317-318, Warner 1993: 137, Roberts & Roussou 2003: 39, Los 2005: 82-88, Fischer & van der Wurff 2006: 149). I also showed here that *gan* and *can*, as well as *beginnen* differed in their use of complementation: a) the verb *gan* was predominantly found with the plain infinitive in ME and

exclusively with this type of complementation in Scots; b) the variant *can* occurred with both the plain infinitive and the *to*-infinitive in NME and in Scots, but the plain infinitive in other varieties; and c) the verb *beginnen* was found with different types of complementation, although the (*for*) *to*-infinitive became the norm for this verb during the ME period, in line with changes that saw the gradual increase in the application of the *to*-infinitive with lexical verbs. I found similar types of complementation occurring with *gan* and *can*, as well as *beginnen* in the NME and ESc data sets, although there were no instances of *can* with the *to*-infinitive in the data sets used in this study. Despite efforts made in the literature to account for the choice of *gan* and *can* with the plain infinitive using metrical and stylistic factors, I argued in this chapter that the choice of this verb and its variant with this kind of complementation in the NME and ESc data sets is the result of this verb and its variant having undergone changes in the underlying structure, i.e. reanalysis, but also analogy. I also showed that reanalysis and analogy could be seen in the case of *beginnen* with its types of infinitive, resulting in instances of *begouth*. As grammaticalization involves morphological, phonological and syntactic change, correlated with semantic change and change in context, only the verb *gan* and *can* show a greater degree of auxiliatation than the verb *beginnen* in the NME and ESc data sets. In the next chapter, the semantics of these verbs will be discussed.

Chapter Five – Changes in Meaning

5.1 Introduction

In this chapter, I am concerned with the semantics of *gan* and *can*, as well as *beginnen*. Unlike the etymologically related *beginnen*, the verb *gan* and its variant *can* usually occur with a non-ingressive meaning in ME and Scots, often described as being similar to that of periphrastic *do*, especially when this verb and its variant are found with the plain infinitive (this also applies to the variant form *couth*, which is exclusively used with non-ingressive meaning) (*DSL* s.vv. *can* v. 2; *gan* p.t., *MED* s.vv. *can* v.; *ginnen* v. 3b, *OED* s.vv. *can* v.² I 1a; *gin* v.¹ 1a, Taylor 1917, Funke 1922: 14, Koziol 1932: 133, Mustanoja 1960: 611, Kerkhof 1966: 159, Visser 1969: 1379, 1571-1575, 1577-1579, Brinton 1981: 159, 161; 1983: 235, 236; 1988: 120, Denison 1993: 322, Ogura 1998: 300). Although Warner (1993) does not identify changes in meaning as one of the criteria for auxiliaryhood, it appears that the occurrence of *gan* and *can* with non-ingressive meaning in ME and Scots is typical of semantic bleaching and so grammaticalization. However, instances of this verb and this variant have also been found with an ingressive meaning in these varieties, typically when the *to*-infinitive is used, but this is not always the case (*DSL* s.v. *gan* p.t., *MED* s.vv. *can* v.; *ginnen* v. 1-3a, *OED* s.vv. *can* v.² I 2; *gin* v.¹ 1a, Eienkel 1891: 89-90, Taylor 1917: 576, 583, 589, Beschorner 1920: 12, Häusermann 1930: 18, Koziol 1932: 131-133, Mustanoja 1960: 613, Kerkhof 1966: 154-156, Visser 1969: 1379-1380, 1573, Brinton 1981: 159; 1983: 235; 1988: 120). Furthermore, in some instances of *gan* and *can*, either a non-ingressive or an ingressive interpretation is possible, with Visser (1969: 1379) arguing that the distinction between such meanings of this verb and its variant is “largely arbitrary” (cf. Eienkel 1891, Taylor 1917: 55, Kerkhof 1966: 158-159, Brinton 1981: 159, Denison 1993: 322). Therefore, it appears that in ME and Scots, the semantics of *gan* and *can* is more complex, and so more interesting, especially when compared to that of the etymologically related *beginnen*.

The information presented here is organised as follows. In Section 5.2, I provide an overview of meanings with which *gan* and *can*, as well as *beginnen* are found in ME

and Scots. I show that: a) *gan* is predominantly found with non-ingressive meaning, with instances of this verb encountered in ingressive sense, too, especially when found with the *to*-infinitive; b) *can* is said to be used either exclusively or predominantly with non-ingressive meaning, whereas the variant form *couth* is found with non-ingressive meaning only; and c) *beginnen* is used with ingressive meaning in ME and Scots (this also applies to the variant form *begouth* in Scots), with isolated instances of this verb being used in a non-ingressive sense in the former variety. I then present my findings of the types of meaning with which *gan* and *can*, as well as *beginnen* are found in the NME and ESc data sets. In Section 5.3, I provide a review of how various sources have dealt with establishing the semantics of *gan* and *can*. This is followed by an analysis in Section 5.4, where I argue that instances of *gan* and *can* characterised by non-ingressiveness in the NME and ESc data sets are the outcome of semantic bleaching, whereas the existence of others, categorised here as ingressive and ambiguous, are characterised by either opacity or persistence. I also show in that section that in a small number of instances of *beginnen* in the NME data set, two interpretations of this verb's meaning are possible. The key to making these discoveries about the classification of meanings of *gan* and *can* is not only to look at such constructions in isolation, or whether this verb and its variant occur with a punctual (achievement) verb in infinitive or adverbials of duration or iteration (in addition to whether *gan* and *can* with a punctual (achievement) verb are used with singular count nouns) (Brinton 1981: 173; 1983: 240-241; 1988: 157-158), but also to consider other contextual information in a text.

5.2 Meaning of *Gan*, *Can* and *Beginnen* in Middle English and Scots

5.2.1 *Gan* and *Can*

Despite being etymologically related to *beginnen*, which conveys the idea of ingression, the verb *gan* and its variant *can*, primarily or in many cases, have a non-ingressive meaning, especially when occurring with the plain infinitive, whereas the variant *couth* is described as exclusively non-ingressive (*DSL* s.vv. *can* v. 2; *gan* p.t., *MED* s.vv. *can* v.; *ginnen* v. 3b, *OED* s.vv. *can* v.² I 1a; *gin* v.¹ 1a, Taylor 1917, Funke 1922: 14, Koziol 1932: 133, Mustanoja 1960: 611, Kerkhof 1966: 156-158,

Visser 1969: 1379, 1571-1575, 1577-1579, Brinton 1981: 159, 161; 1983: 235, 236; 1988: 120, Denison 1993: 322, Ogura 1998: 300). This follows from claims in the literature that the more grammaticalized variant takes the plain infinitive, as shown to be the case with verbs characterised by auxiliation (modal and auxiliary verbs) in ME (Warner 1993: 103). However, even though the non-ingressive meaning is usually associated with *gan* found with the plain infinitive, some sources claim that this kind of meaning could also be found in instances of this verb with the *to*-infinitive in ME and Scots. For example, Warner (1993: 139) argues that in ME, “[in] meaningless ‘auxiliary’ uses ... even here there is variation between the *to*-infinitive and the plain infinitive [with *gan*]”, whereas Moessner (1997: 114) emphasises the fact that in OSc, this construction has a different grammatical status from modal (auxiliary) verbs, “[a]n expansion by *gin* (present) or *gan* (past) can be followed by the bare infinitive or by the *to*-infinitive. Although formally similar to the expansion by a modal auxiliary, it is syntactically and semantically different; syntactically, because it does not combine with other expansions, semantically, because *gin/gan* does not add a modal component.”

Given statements in the literature that *gan* and *can* usually occur with non-ingressive meaning in ME and Scots, especially when found with the plain infinitive, this non-ingressive meaning has been described to be similar to that of periphrastic *do* (*DSL* s.vv. *can* v. 2; *gan* p.t., *MED* s.vv. *can* v.; *ginnen* v. 3b, *OED* s.vv. *can* v.² I 1a; *gin* v.¹, Funke 1922: 26, Mossé 1952: 103, Mustanoja 1960: 602, Visser 1969: 1572, 1577, Traugott 1972: 141, Brinton 1996: 67, Ogura 1997: 425, Sims 2008: 151), or, simply, as semantically empty (Funke 1922: 24-25, Smyser 1967, Sims 2008: 15, 136, 151, 171; 2014). As this highlights the difficulty with the categorisation of meaning of this verb and its variant, also suggesting that *gan* and *can* might have been used interchangeably with *do*, attempts have been made to describe this non-ingressive meaning in terms of aspectual terminology, more specifically the term ‘perfectivity’.⁸⁵ This term is related to grammatical aspect, i.e. the possibility of

⁸⁵ The notion ‘aspect’ deals with the temporal structure of a clause in terms of the time intervals (i.e. phases) formed in the interpretation of the event or situation denoted by that clause. Sasse (2006: 31) argues, “[t]he fundamental criterion is the inclusion or non[-]inclusion of starting points and/or end-points (‘boundaries’) in the conceptualization of the [event or] situation.” In semantic view, an

using a verb form to denote meanings linked to the internal temporal structure of an event or situation. Perfective aspect is, therefore, seen as the actualisation of an event or situation as a whole, in its entirety, without references to its temporal structure, i.e. its beginning, middle and end. This highlights a difference between an event or situation that is portrayed as complete, i.e. whole (telic), and one that is presented as completed, i.e. finished (atelic) (Declerck, Reed & Cappelle 2006: 28-31). Therefore, Los (2000: 260-261; 2005: 305) claims that both OE *onginnan* and OE *beginnan* had lost their ingressive meaning and “were bleached to something like a perfective auxiliary” when found with the plain infinitive in Ælfric (cf. Funke 1922: 5-6, Mustanoja 1960: 610-611, Richardson 1991a: 34-59 on OE *onginnan*), while Funke (1922: 24-25) and Brinton (1996: 76) use this kind of terminology to describe the meaning of ME *gan*, “[g]an is active and perfective, and the accompanying infinitive denotes a dynamic, telic event, which carries the narrative forward”.

However, in addition to the non-ingressive meaning, both *gan* and *can* have been described as having an ingressive meaning in some contexts in ME and Scots (*DSL* s.v. *gan* p.t., *MED* s.vv. *can* v.; *ginnen* v. 1-3a, *OED* s.vv. *can* v.² I 2; *gin* v.¹ 1a, Einkenel 1891: 89-90, Taylor 1917: 576, 583, 589, Beschorner 1920: 12, Häusermann 1930: 18, Koziol 1932: 131-133, Mustanoja 1960: 613, Kerkhof 1966: 154-156, Visser 1969: 1379-1380, 1573, Brinton 1981: 159; 1983: 235; 1988: 120). This type of meaning is usually found when this verb and its variant occur with the *to*-infinitive. However, this is not always the case. For example, some sources like the *DSL* (s.v. *gan* p.t.) resort to listing “[b]egan, did” for *gan* in Scots, which suggests that this verb can indeed have either a non-ingressive or ingressive meaning in this variety. But, some clues as to when to apply the meaning of ‘began’ and ‘did’ are missing from some dictionary descriptions (cf. *MED* s.vv. *can* v.; *ginnen* v., *OED*

aspectual situation can be seen as either completed/perfective (an external viewpoint) or on-going/imperfective (an internal viewpoint). In formal terms, aspect is expressed in a grammatical way, for example, by means of aspectual grammatical constructions in ModE (Boogaart 2004: 1165-1166). Brinton (1988: 3) observes that aspect is “a matter of the speaker’s viewpoint or perspective on a situation”, resulting in the speaker portraying “as completed (perfective aspect), or as ongoing (imperfective aspect), or as beginning (ingressive aspect), continuing (continuative aspect), ending (egressive aspect), or repeating (iterative or habitual aspect)”. Due to the spatial constraints of this study, it is not possible to present here a literature review on aspect. Cf. Comrie (1976), Dowty (1979), Brinton (1981, 1988, 1996), Saurer (1984), Bybee (1985), Dahl (1985), Binnick (1991), Hataav (1993), Tobin (1993), Mellor (1995), C. S. Smith (1991; 1997) and Kabakčiev (2000) amongst others for further details.

s.v. *gin* v.¹). Given the difficulty with separating ingressive from non-ingressive meanings of *gan* and *can* in ME and Scots and distinctions between such instances of this verb and its variant being “‘largely arbitrary’” (Visser 1969: 1379) (cf. Einkenkel 1891, Taylor 1917: 55, Kerkhof 1966: 158-159, Brinton 1981: 159, Denison 1993: 322), Visser (1969: 1573) speculates about what he calls ‘typical’, i.e. non-ingressive, instances of *gan* with the plain infinitive, “[n]ot in all of them is the ingressive meaning of *gan* = *commence* demonstrably non-existent”. By doing so, he probably refers to a situation characterised by the non-final stages of the process of grammaticalization where traces of the original meaning of a construction still persist (Hopper 1991, Hopper & Traugott [1993] 2003). Furthermore, Visser (1969: 1575, 1578) makes similar statements about borderline cases of *gan* with the *to*-infinitive and *can* with the plain infinitive, this once again highlighting that there might, indeed, be some difficulty in sieving non-ingressive instances from ingressive ones of these verbs.

Although both non-ingressive and ingressive meanings are ascribed to *gan* and *can* by the sources referred to above, the verb *ginnen* in absolute, transitive and intransitive uses (finite and non-finite forms; followed by infinitive, with or without *to*) is typically used with ingressive meaning (*MED* s.v. *ginnen* v. 1-3a, *OED* s.v. *gin* v.¹ 1a, 2-3, Beschorner 1920: 12, Funke 1922, Häusermann 1930: 18, Koziol 1932: 131-133, Mustanoja 1960: 612, Kerkhof 1966: 154-156, Visser 1969: 1379-1380, Brinton 1981: 159; 1988: 120; 1996: 68). This occurrence of *gan* and *can* (third person singular and in the preterite) with the plain infinitive in ME and Scots, whereas the presence of *ginnen* (in the present indicative) as an independent verb and with infinitive in ME may be indicative that a split has occurred: *gan* and *can* undergo changes in meaning in ME and Scots, whereas *ginnen* continues to be exclusively used with the ingressive meaning in these varieties. However, the *MED* (s.v. *ginnen* v. 3b) also claims that instances of *ginnen* (in the present indicative) can be found in auxiliary use, “[a]s a weak auxiliary used with infinitives to form phrases denoting actions or events as occurring (rather than as beginning to occur) ... with present and future meaning *ginneth maken*, doth make, makes”, with examples from

SME.⁸⁶ Overall, therefore, it appears that *gan* and *can* have either non-ingressive or ingressive meanings in ME and Scots, although this distinction is not always clear-cut; it also seems that the non-ingressive meaning is preferred when this verb and its variant are used with the plain infinitive. In contrast, the verb *ginnen* always occurs with the ingressive meaning when found functioning as an independent verb or with infinitive. Let us now turn to the types of meaning of *gan* and *can* in the NME and ESc data sets to see whether the situation in these data sets with regards to the meaning of this verb and its variant resembles that described above.

The types of meaning associated with *gan* and *can* in the NME and ESc data sets are provided in Table 45, Table 46 and Table 47 (cf. Section 2.3.1 on the semantic criteria used in the selection process of instances of *gan* and *can* in this study). This division into different categories of meanings appears to depend on the classification

⁸⁶ The *MED* (s.v. *ginnen* v. 3b b) lists one instance of *ginnen* with the plain infinitive from *Cursor Mundi* (Edinburgh, Royal College of Physicians MS), as in (1) below, whose dialect is described as “WRY [West Riding of Yorkshire]” (*eLALME* [formerly LP 14, LP375]). However, although this example is listed under the sense of ‘[a]s a weak auxiliary used with infinitives to form phrases denoting actions or events as occurring (rather than as beginning to occur): do, did’, *gin* occurs as *you* or its variant forms in other MSS of this work, as in (2) below. Furthermore, in the Edinburgh, Royal College of Physicians MS, <g> is used for both /g/ and /j/, as in (3) below. I would, therefore, argue that *gin* in (1) below is a spelling variant of ‘you’, rather than an instance of *ginnen*.

- (1) *Ful wel mai I Of þe prophete*
 Full well know I of the prophet
gin telle
 you tell

(a1400 *Cursor Mundi* 18993-18994; *MED* s.v. *ginnen* v. 3b, cf. Morris 1874-1893: 1587)

‘I know very well of the prophet you speak about’

- (2) *O þe prophete yow tell dau* (Cotton Vespasian A.3 MS; Morris 1874-1893: 1088; hand A, covering lines 18,029-20,064, in a dialect of “YWR [Yorkshire, West Riding]” *eLALME* LP18)
Tell you of the prophete dauy (Fairfax – Arundel 57 and Laud 416 MSS; Morris 1874-1893: 1088)
ȝu tell of þe prophete dauy (Göttingen MS Theol. 107; Morris 1874-1893: 1089; dialect of “LIN/YWR [Lincolnshire/Yorkshire, West Riding]” *eLALME*)
Telle ȝou of þe prophete dauy (Trinity College, Cambridge MS R.3.8; Morris 1874-1893: 1089; dialect of “STF [Staffordshire]” *eLALME* LP36)

- (3) *Giu sal forgiuig be giur sak*
 You shall forgiven by your sake
 (*Cursor Mundi* 19019; Morris 1874-189: 1587)

‘You shall [be] forgiven by your sake’

of the verb in infinitive in accordance with its Aktionsart properties, but also other elements within a clause, as will be discussed in Section 5.3, as well as other contextual information. The information contained in these tables, therefore, shows that the verb **gan** is used with a non-ingressive and ingressive meaning in the NME data set in Period I and Period II, as in (1) below and (2) below; but, there are also cases where the meaning of this verb cannot be precisely determined in that it can be seen as either ingressive or non-ingressive, depending on one's interpretation of the Aktionsart of the verb in infinitive, as in (3) below, with contextual clues sometimes missing, as will be discussed in Section 5.4.

Table 45

*Types of Meaning of **gan** with Infinitive in the NME Data Set in Period I and Period II*

	<i>gan</i>					
	Non-ingressive		Ingressive		Ambiguous	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Period I	5	29	1	6	11	65
Period II	40	53	6	8	29	39

Notes: *N* – Observed absolute frequency
 % – Per-cent of **gan** with its respective meaning

Table 46

*Types of Meaning of **can** with Infinitive in the NME Data Set in Period I and Period II*

	<i>can</i>			
	Non-ingressive		Ambiguous	
	<i>N</i>	%	<i>N</i>	%
Period I	4	50	4	50
Period II	8	67	4	33

Notes: *N* – Observed absolute frequency
 % – Per-cent of **can** with its respective meaning

Table 47

*Types of Meaning of **can** with Infinitive in the ESc Data Set in Period I*

	<i>can</i>			
	Non-ingressive		Ambiguous	
	<i>N</i>	%	<i>N</i>	%
Period I	5	63	3	38

Notes: *N* – Observed absolute frequency
 % – Per-cent of **can** with its respective meaning

- (1) *Clement to the mynstralles gan go*
Clement to the musicians did go
And gafē some a stroke and some two
And gave some a stroke and some two
 (1420-1500 *Octavian* l. 1205-1206; *TEAMS*)

‘Clement went to the musicians and gave some a blow and others two’

- (2) *When scho was flemyd that was so gent,*
When she was banished that was so gentle,
Ilke a lorde to hys lande es went,
Each a lord to his land is gone,

For sorow thaire hertes gan blede.
 For sorrow their hearts began to bleed.
 (1420-1500 *Octavian* l. 289-291; *TEAMS*)

‘When she, who was so gentle, was banished, each lord went to their country, and their hearts began to bleed with sorrow.’

(3) *Therewith the lady bygan to wake,*
 Therewith the lady did/began to wake,
A dolefull gronyng gan scho make
 A distressing groaning did/began to she make
 (1420-1500 *Octavian* l. 172-173; *TEAMS*)

‘With that the lady did/began to wake, a distressing groan she did/began to make’

In contrast to *gan*, the variant *can* appears to be recorded with a non-ingressive meaning in the NME and ESc data sets, as in (4) below and (5) below, respectively, with no instances appearing to have ingressive meaning. However, like with *gan*, there are also instances of the above variant, where its meaning can be interpreted in two ways, as either ingressive or non-ingressive, as in (6) below and (7) below. This is in line with statements in the literature, where attention is drawn to the problem of the meaning of *gan* and *can* not always being overtly non-ingressive in ME (Einenkel 1891, Taylor 1917: 55, Kerkhof 1966: 158-159, Visser 1969: 1379, Brinton 1981: 159, Denison 1993: 322), with Visser (1969) specifically pointing out that the distinctions between the meanings of such variants are arbitrary. Overall, therefore, while the situation of having different types of meaning of *gan* and *can* in the NME and ESc data sets demonstrates that there are changes taking place in relation to the meaning of this verb and its variant, it also shows that such different meanings can be the result of the difficulty with interpretation of the meaning of this verb and its variant, as already referred to in this section, but fully addressed in Section 5.3.

(4) *Florent to the stede can gone;*
 Florent to the place did go;
So feyre an hors sye he never none
 So fair a horse saw he never none
 (1420-1500 *Octavian* l. 724-725; *TEAMS*)

‘Florent went to the place; never before did he see such a beautiful horse’

- (5) *Sad to hyme agane, “Thu ga*
Said to him again, “Thou go
Til hyre that can this demand ma
To hir that did this demand make
And spere at hyre grathly.
And ask at her promptly.
(1375-1420 *St. Andrew and the Three Questions* l. 1084-1087; *TEAMS*)

‘[the porter] said to him again, “Go to her that made this demand and ask her immediately.”’

- (6) *And gave hym corne and haye,*
And gave him corn and hay,
And sethyn he can hym kembe and dyght
And then he did/began to him comb and groom
(1420-1500 *Octavian* l. 744-745; *TEAMS*)

‘And gave him corn and hay, and then he did/began to comb and groom him’

- (7) *Thane watir one thame men can caste,*
Then water on those men did/began to cast,
And thai ourecome at the last.
And they overcame at the last.
(1375-1420 *The Life of St. Julian Hospitaller* l. 346-348; *TEAMS*)

‘Then [she] did/began to throw water on those men, and they finally recovered.’

The data in the tables also shows that instances of **gan** and **can** with non-ingressive meaning are the largest group in the NME data set, but in Period II only. Such instances form 53% and 67% of all examples of this verb and its variant there, respectively. However, a different situation can be observed for this verb and its variant in Period I. When comparing percentages of different types of meaning of instances of **gan** in Period I, it appears that the highest percentage of the type of meaning of this verb is from the ‘ambiguous’ category, whereas for the variant **can** in the same period, the two types of meaning, namely non-ingressive and ambiguous, are equally distributed there. Let us apply the log-likelihood statistical test to see if the differences in the frequency of **gan** and **can** between periods in this data set are

statistically significant. So, the null hypothesis is that there is no correlation between the differences in the frequencies of the categories of meaning of *gan* and *can* over the periods in the NME data set, with the alternative hypothesis that there is such a correlation. In relation to *gan*, the log-likelihood value for the category of non-ingressive meaning of this verb is 7.48, the difference in frequency for this category over Period I and Period II is statistically significant at p being 0.01 (99th percentile, 1% level), meaning that we should reject the null hypothesis; with reference to the log-likelihood values for the categories of ingressive and ambiguous being 0.73 and 0, respectively, this shows that we should accept the null hypothesis that there is no correlation between the differences between the frequencies of *gan* in these two categories over periods at p being 0.05 (95th percentile, 5% level). With respect to *can*, the log-likelihood values for the categories of non-ingressive and ambiguous meaning identified for this variant are 0.18 and 1.75, respectively, also suggesting no correlation between the types of meaning associated with this variant over periods at the p value of 0.05 (95th percentile, 5% level). So, the probability of the result – i.e. the difference between the two periods in the NME data set – happening by chance is 1% for *gan* in the non-ingressive sense and 5% for this verb in ingressive and ambiguous meanings, as well as the variant *can* in the non-ingressive and ambiguous meanings.

In relation to in the ESc data set, only two categories of meaning can be identified for *can*, namely non-ingressive, which contains the highest percentage of instances of this variant there (63%), and ambiguous (38%). When comparing the percentages of the types of meaning of this variant with those of *gan* and *can* in the NME data set in Period I, it appears that there is a tendency for *can* in the NME and ESc data sets to be at the forefront in terms of its occurrence with the non-ingressive meaning. However, this is affected by the low number of instances of this variant in both data sets – should the absolute (raw) frequency of this variant be higher in both data sets, the percentage of non-ingressive meaning would proportionally change. Therefore, I test these observations against the null hypothesis that there is no correlation between the differences in the frequencies of the categories of meaning of the variant *can* over regions, i.e. NME and ESc varieties represented by textual evidence in the NME

and ESc data sets, respectively. So, with the log-likelihood being 4.90 for the category of the non-ingressive meaning of the variant *can* in these data sets, there is no correlation between the frequencies of the categories of meaning of this variant over regions at the p value of 0.05 for the non-ingressive meaning (95th percentile, 5% level). However, with the log-likelihood value 3.72 for the category of ambiguous of this variant, and 0.05 for the ambiguous one, the null hypothesis should be accepted (95th percentile, 5% level). Taking these together, the probability of the result – i.e. the difference between the two regions – happening by chance is 1% for *can* in the non-ingressive sense and 5% for this variant characterised by the ambiguous meaning.

5.2.2 *Beginnen*

Unlike the verb *gan* and its variant *can*, the verb *beginnen* in its aspectual function almost exclusively occurs with an ingressive meaning in ME (*MED* s.v. *biginnen* v. 4, *OED* s.v. *begin* v. ¹, Einkenel 1891: 90, Funke 1922: 11-12, Häusermann 1930: 21-22, Koziol 1932: 133-134, Mustanoja 1960: 611, Visser 1969: 1373, Brinton 1981: 181, 186; 1983: 241; 1988: 159, Ogura 1997: 403-428, Sims 2008: 151-152; 2014), and solely with this type of meaning in Scots; this also applies to the variant form *begouth* (*DSL* s.v. *begin* v. 1).⁸⁷ However, in ME, this verb can also be found with a non-ingressive meaning (Los 2000; 2005 finds instances of OE *beginnan* with bleached meaning when found with the plain infinitive in Ælfric). For example, the *MED* (s.v. *biginnen* v. 6) argues about this verb, “emphasizing the occurrence of an event or act denoted by the following infin.: e.g. **beginneth springe**, does leap, **began to calle**, did call [Cp. the frequent use of **ginnen**, **gan** with this function.]”, as in (8) below (cf. Häusermann 1930: 21-22, Brinton 1981: 185-186). Instances of this verb with the non-ingressive meaning are recorded in 1325 *The Flemish Insurrection*, c1375 Chaucer’s *Canterbury Tales*, a1400 *Cursor Mundi*, a1425 Chaucer’s *Troilus and Criseyde*, c1430 Chaucer’s *Legend of Good Women*, a1470 Malory’s *Works* and a1500 *Generides*, therefore written in SME and NME, as well as in EModE (*MED* s.v. *biginnen* v. 6). By comparing the non-ingressive meaning of

⁸⁷ Sources like Reeves (1893), Beschorner (1920), Romaine (1981), Moessner (1989), Macafee (1992/93), Meurman-Solin (1993a; 1993b; 1993c), King (1997), Moessner (1997), Jumpertz-Schwab (1998) and Bugaj (2004) make no special reference to *beginnen* in Scots.

this verb to that of *gan*, this source also implies that *beginnen* might be subject to a further degree of grammaticalization. But, the occurrence of *beginnen* with the non-ingressive meaning also only appears to concern isolated instances of this verb, with a general consensus in the literature that overall, the verb *beginnen* mainly expresses an ingressive meaning in ME and Scots (Brinton 1981: 181, 186, Sims 2008: 151-152, cf. Beschorner 1920: 15, Mustanoja 1960: 612).

(8) *Miht he noght blinne To bihold þat fair maydene,*
 Might he not stop to behold that fair girl,
Hou all hir dedis begane hir seme
 How all her deeds did her befit
 (a1400 *Cursor Mundi* 3311 (Göttingen MS Theol. 107); *MED* s.v. *biginnen* v. 6)

‘He cannot cease to behold that fair maiden, how all her deeds befitted her’

The frequencies of occurrence of different types of meaning found with the verb *beginnen* in the NME and ESc data sets are given in Table 48 and Table 49, respectively.

Table 48

*Types of Meaning of **beginnen** as an Independent Verb and with Infinitive in the NME Data Set in Period I and Period II*

	<i>beginnen</i>					
	Other		Inf		Inf	
	Ingressive		Ingressive		Ambiguous	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Period I	24	100	27	100	-	-
Period II	38	100	131	97	4	3

Notes: *N* – Observed absolute frequency
 % – Per-cent of *beginnen* with its respective meaning

Table 49

*Types of Meaning of **beginnen** as an Independent Verb and with Infinitive in the ESc Data Set in Period I and Period II*

	<i>beginnen</i>			
	Other		Inf	
	Ingressive	Ingressive	Ingressive	Ingressive
	<i>N</i>	%	<i>N</i>	%
Period I	7	100	1	100
Period II	55	100	11	100

Notes: *N* – Observed absolute frequency

% – Per-cent of **beginnen** with its respective meaning

According to the above tables, **beginnen** is exclusively used in the ingressive sense when found as an independent verb in these data sets in both periods, as in (9) below and (10) below, respectively. This ingressive meaning also characterises all instances of this verb with infinitive in the NME data set in Period I, as in (11) below, and in the ESc data set in Period I and Period II, as in (12) below. This also includes instances of the variant form **begouth** in the ESc data set. As the verb **beginnen** continues to be used with the ingressive meaning in the NME and ESc data sets, my findings are similar to those found in the literature, as discussed above. However, the difference is that some of those sources mentioned above see some instances of this verb as non-ingressive (*MED* s.v. *biginnen* v. 6, Häusermann 1930: 21-22, Visser 1969: 1580-1581, Brinton 1981: 185-186), except for Visser (1969: 1581), who acknowledges the fact that sometimes, the meaning of **beginnen** may be categorised as ambiguous, too. Furthermore, although the greater majority of instances of **beginnen** with infinitive in the NME data set in Period II (97%) are used with the ingressive meaning, there are a small number of instances of this verb (3%) characterised by a degree of ambiguity in relation to meaning, resulting in two possible interpretations, namely either non-ingressive or ingressive, as in (13) below. Taking into account the meanings of **gan** and **can** as well as that of **beginnen** in the NME and ESc data sets, it is evident that a split has taken place, with **gan** and **can**

being predominantly used in a non-ingressive sense in these data sets, whereas *beginnen* being used with an ingressive meaning there.

- (9) *Pe first: is writen bifore, & bigynnes at*
 The first: is written before, & begins at
Ante oracionem prepara a[nimam] t[uam] &
 Ante oracionem prepara a[nimam] t[uam] &
lastis h[id]er
 lasts hider
 (1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 142; *ICoMEP*)

‘The first is written before, and begins at *Ante oracionem prepara animam tuam* [before prayer prepare thy soul] and continues here’

- (10) *HER BEGYnNYS A NOBYL TRETYSE MADE*
 Here begins a noble treatise made
OF A GUD PHEsICIAN jOHN OF BURDOUSE
 of a good physician John of Burdouse
FOR MEDICENE AGAYNE YE PESTILENS jWYLL
 for medicine against the pestilence evil
 1420 *Liber S. Marie de Calchou* 164; *LAOS*)

‘Here begins a great book written by a good physician John of Burdouse on medicine against the evil’

- (11) *Pe heued biginnes for to scak,*
 The head begins for to shake,
His hend vnquemli for to quak,
 His hand unsettlingly for to tremble,
It crepes crouland in his bak
 It creeps crawling on his back
 (1350-1420 *Cursor Mundi* p. 212; *HCET*)

‘The head begins to shake, and his hand [begins] to unsettlingly tremble; it creeps crawling on his back’

- (12) *And than the wellis begynnys till habound jn*
 And then the wells begins to overflow in
wateris and stryndis to spring
 waters and streams to spring
 (1420-1500 *The Prose Works of Sir Gilbert Hay* p. 98; *HCOS*)

‘And then the wells begin overflowing and streams [begin] to flow [out of the ground]’

- (13) *Therewith the lady bygan to wake,*
 Therewith the lady did/began to wake,
 A dolefull gromyng gan scho make
 A distressing groaning did/began to she make
 (1420-1500 *Octavian* l. 172-173; *TEAMS*)

‘With that the lady did/began to wake, a distressing groan she did/began to make’

Let us now turn to the statistical analysis of the meanings associated with the verb *beginnen*. With the null hypothesis being that there is no correlation between the differences in the frequencies corresponding to the meanings with which *beginnen* is found over periods in the NME data set, the log-likelihood values are as follows: a) with reference to the ingressive meaning when *beginnen* functions as an independent verb, this value is 3.43; this shows that we should accept the null hypothesis at $p < 0.05$ (95th percentile, 5% level); b) in relation to the ingressive meaning when this verb occurs with infinitive, this value is 9.99, meaning that the null hypothesis should be rejected at $p < 0.01$ (99th percentile, 1% level); and c) for the non-ingressive meaning when this verb is found with infinitive, the log-likelihood is 2.61, meaning that the null hypothesis should be accepted at $p < 0.05$ (95th percentile, 5% level). Therefore, only the differences in the frequencies of *beginnen* in the ingressive sense functioning as an independent verb and in the non-ingressive when this verb is found with infinitive indicate the probability of happening by chance between Period I and Period II in the NME data set, with there being a correlation in relation to the differences in the frequencies of this verb in the ingressive meaning and this verb functioning with infinitive between both periods and in this data set. With reference to the verb *beginnen* in the ESc data set, with the null hypothesis being similar, the log-likelihood value is 1.81 for this verb functioning as an independent verb and 0.05 for this verb with infinitive ($p < 0.05$, 95th percentile, 5% level), indicating strong evidence against the null hypothesis. This means that we have to accept the alternative hypothesis, namely that there is a correlation in the frequency of *beginnen* with ingressive meaning over both periods in the ESc data set.

Overall, therefore, in line with statements in the literature, the findings on the different categories of meaning of *gan* and *can* in the NME and ESc data sets presented here indicate that, primarily, the occurrence of this verb and its variant is with a non-ingressive meaning, although there are a number of instances of this verb and its variant in the categories of ambiguous and ingressive, too. The verb *beginnen*, on the other hand, is found with the ingressive meaning in these data sets, except for a small number of instances of this verb with a non-ingressive meaning in the NME data set. In Chapter Four, I indicated that *gan* and *can* undergoes a number of analogical changes, one of which was the co-occurrence of this verb and its variant with punctual (achievement) verbs in infinitive, which are incompatible with ingression. Furthermore, it has also been observed that this verb and its variant can co-occur with adverbials of duration or iteration, this occurrence resulting in a non-ingressive interpretation of *gan* and *can*. However, an ingressive interpretation is likely when *gan* and *can* with a punctual (achievement) verb in infinitive is found with plural count nouns and mass nouns) (Brinton 1981: 173; 1983: 240-241; 1988: 157-158). In the subsequent section, I discuss the proposals provided by various sources to help determine the meaning of this verb and its variant in these varieties.

5.3 Determining the Meaning of *Gan*, *Can* and *Beginnen*

It has been observed in the literature that the non-ingressive meaning of *gan* and *can* in ME and Scots, as well as that of the etymologically related OE *onginnan* and OE and eME *aginnan*, is the result of the Aktionsart of the verb in infinitive found with the above verbs, but also other elements in a clause (Wuth 1915: 55, Beschorner 1920: 13-14, Funke 1922: 12, 24, Kerkhof 1966: 156-157, Brinton 1981: 169-174; 1983: 239-241; 1988: 154-161; 1996: 75-78). In semantic terms, Aktionsart is used to refer to a typology of a state of affairs, i.e. it deals with the characteristics of a verb describing an action, event or state. The traditional view stipulates that Aktionsart is concerned with the inherent temporal properties, i.e. “the phasal time structure of the event it denotes”, such as the intervals of the development of an event, namely beginning, end, repetition (Sasse 2006: 31). However, in formal terms, Aktionsart takes into account the way in which aspectual notions are expressed in a language (Boogaart 2004: 1166-1167). This is why different quantificational or

determinational characteristics of other elements, for example NPs and PPs, can result in different Aktionsart interpretations, therefore being independent of the verb's 'basic' Aktionsart (Sasse 2006: 31). When discussing the meaning of *gan* and *can* (whilst making references to that of *beginnen*, too), I distinguish between the 'basic' Aktionsart of the verb and the other elements contributing to the Aktionsart interpretation. Nevertheless, this is not to suggest that they are not interconnected. On the contrary, I argue that they are linked, and so each needs to be taken into account when explaining the aspectual impact of the 'basic' Aktionsart of the verb in infinitive and other elements in a clause on the meaning of *gan* and *can* in ME and Scots, including the NME and ESc data sets.⁸⁸

However, before reviewing the literature on approaches to the study of the meaning of *gan* and *can*, it is necessary to introduce the idea behind verb typologies associated with 'basic' Aktionsart of the verb. Out of all typologies, Vendler's (1957; 1967) typology is the most widely used in the aspectual literature, including studies dealing with the semantics of *gan* and *can* – this typology offers a four-way classification of verbs into 'states', 'activities', 'accomplishments' and 'achievements', although a category of 'semelfactives' has been added by Comrie (1976). Other typologies subdivide on the basis of classes or categories by Vendler (1957; 1967), but with some variation: for example, Carlson (1981) and Bach (1986) offer more classes, whereas Kenny (1963), Mourelatos (1978) and Dowty (1979) offer fewer. However, since the number of classes or categories is not relevant to this study, an overview of such typologies will not be presented here. Instead, while analysing instances of *gan* and *can* in the NME and ESc data sets, I use Vendler's (1957; 1967) and Comrie's (1976) classifications here as a point of reference, even though they refer to the 'basic' Aktionsart of the verb. Furthermore, I make references in the remainder of this chapter to descriptions of the classes or categories of the 'basic' Aktionsart provided by Dowty (1979), C. S. Smith (1991) and Piñón (1997). Finally, since the category of achievement verbs is of particular interest here

⁸⁸ Due to the spatial constraints of this study, it is not possible to present here a literature review on the Aktionsart, but only on the effects of it on *gan* and *can*. Cf. also Vendler (1957; 1967), Kenny (1963), Leech (1971), Verkuyl (1972), Comrie (1976), Mourelatos (1978), Dowty (1979), Dahl (1985), Brinton (1988), Olsen (1994), Tenny (1994), C. S. Smith (1991; 1997), Bhat (1999) and Kabakčiev (2000) amongst others on information on the Aktionsart and its effects.

because such verbs are punctual and so incompatible with ingressive aspect, but also found with some instances of *gan* and *can* in ME and Scots, I mainly concentrate on this group here. Let us turn to a description of achievement verbs.

Vendler (1967: 24-26) defines achievements as “occur[ing] at a single moment”, “involv[ing] unique and definite time instants”, “[not being] regarded as voluntary (or involuntary) actions”, adding that “in cases of pure achievement terms the present tense is almost exclusively used as historic present or as indicating immediate future”; an example of an achievement verb found with *gan* is in (14) below.

Achievements, therefore, behave differently from both activities and states in that they are telic, i.e. they indicate that an event or situation is complete in some sense. They do not generally occur with the *for*-phrase, but the *in*-phrase, as in (15) below, time adverbials, and constructions with *spend-an-hour/take-an-hour*, as these “do not measure the smallest interval *during which* the described eventuality takes place” (Piñón 1997: 278). Furthermore, they are not found as complements of verbs like *finish*, i.e. verbs focussing on the moment of inception or cessation of an action or event; in this respect, they would not be compatible with the verb *begin*. Finally, they are also anomalous with (agent-orientated) adverbs like *attentively*, *studiously*, *obediently* and *carefully*, as in (16) below, given that “the referent of the subject NP participated intentionally in the described eventuality” (Piñón 1997: 280), manner adverbs like *quickly* and *slowly*, as in (17) below, but compatible with the adverb *almost* (cf. Dowty 1979: 58-59, C. S. Smith 1991, Piñón 1997). Overall, therefore, achievement verbs are not compatible with expressions of duration, but are compatible with an expression of completion (C. S. Smith 1997: 327).

- (14) *And ryght anon the wympel gan she*
And right immediately the wimple did she
finde
 find

(Chaucer *The Legends of the Good Women* 819; Beschorner 1920: 14, as cited in Brinton 1983: 240; translation in the original)

‘and right away she found (NOT began to find) the wimple’

- (15) *Rebecca reached the summit in five hours* (in a split second) (Piñón 1997: 278)

- (16) #*Rebecca intentionally (attentively, consciously, studiously, vigilantly) reached the summit* (Piñón 1997: 280)
- (17) #*Rebecca quickly (slowly) reached the summit* (Piñón 1997: 280)

I also provide a description of accomplishment and semelfactive verbs for comparison, because of their similarities to achievements. According to Vendler (1967: 26), accomplishments “imply the notion of unique and definite time periods”; in this respect, this category is similar to achievements, but different from them in that it also encompasses an idea of duration and approaching an endpoint incrementally, as in (18) below, i.e. ‘to come out of the state of sleep or unconsciousness; to be roused from sleep, cease to sleep’ (*OED* s.v. *wake* v. II). Semelfactives, like achievements, denote an instantaneous, i.e. punctual, event; however, in contrast to achievements, this event is described as atelic, i.e. presented as being incomplete (C. S. Smith 1997: 55). Therefore, verbs belonging to this category have no preliminary and resultant stages, while their single stages have simultaneous initial and final points, as is the case for the ModE verbs *knock* and *cough*. Furthermore, semelfactives are not compatible with imperfective aspect and durative adverbials, but C. S. Smith (1997: 56-57) argues that this is not always the case – semelfactives frequently undergo reinterpretation, namely to a derived multi-event activity. For example, in relation to imperfective aspect, the focus of it is on an interval linked to an event, but this kind of interval is absent from verbs classified as semelfactives; what is more, even though imperfective aspect can be associated with instantaneous events like *she is reaching the top*, this is not possible for semelfactives, as this type of event has no preliminary stages. In such cases, semelfactives are reinterpreted as a repetitive activity. In relation to durative adverbials, since *John coughed for five minutes* has a durative adverbial, only repetitive information is allowed here.⁸⁹

- (18) *his heorte gon to wakien*
 his heart began to wake
 (c1275 (?a1200) *Lazamon Brut* 9879; Visser 1969: 1379, as cited in Brinton 1983: 240; translation in the original)

⁸⁹ This is different from *John coughed in five minutes*, in which case the adverbial *in five minutes* signals a temporal location, i.e. John coughed after five minutes of something else; with manner adverbials implying duration, semelfactive constructions focus on the beginning of an event or action; for example, *John slowly knocked at the door* is reinterpreted as *John was slow to knock*.

‘his heart began to awake’

However, distinctions into ‘basic’ Aktionsart categories can be ‘slippery’, as there is a problem with the verb classification itself in that some verbs can be put into more than one category. For example, the verb *think* can be considered as either a state verb, as in *I think it’s a great idea*, or an activity verb, as in *I’m thinking about my exam*, whereas the verb *arrive* can be viewed as an activity, i.e. the process of a train appearing at a train station, or an achievement verb, i.e. the moment of reaching one’s destination. Furthermore, with attempts by Vendler (1957; 1967) to classify verbs into aspect categories, such classifications do not take into account the contribution of other elements in a clause. For example, the presence of the NP *a thousand men* found with the verb *kill* results in an iterative interpretation of the event presented here, as opposed to that of *kill the man*, which offers a punctual reading. In this respect, the verb *kill* can be classified as both an accomplishment verb and an achievement verb. Other sources (Dowty 1979, Dorr 1993) propose tests for verb class membership, attempting to formalise and classify various types of lexical and semantic differences using different parameters. However, this highlights problems with applying such tests, due to the fact that context and other constituents within a clause, such as the above-mentioned argumentation, but also adjuncts, often affect items under investigation, resulting in changes of such categories of verbs. For example, the verb *run* can be seen as an activity, as in *Adam ran*, or an accomplishment, as in *Adam ran for a mile* or *Adam ran to the supermarket (= Adam went to the supermarket)*, given the presence of PPs imposing the restriction on the duration of the event or situation presented here.

As stated at the beginning of this section, the effect of the ‘basic’ Aktionsart of the verb in infinitive on the meaning of *gan* and *can* has already been noted in the literature, with other elements within a clause also affecting the meaning of this verb and its variant (Wuth 1915: 55, Beschorner 1920: 13-14, Funke 1922: 12, 24, Kerkhof 1966: 156-157, Brinton 1981: 169-174; 1983: 239-241; 1988: 154-161; 1996: 75-78). However, not all accounts dealing with the meaning of *gan* and *can* note the potential effect of the verb in infinitive on the meaning of this verb and its

variant. Instead, they only provide descriptions of what is being observed, although still making attempts at some sort of classification of the verb in infinitive into different categories. For example, Mustanoja (1960: 610) notes that in Chaucer, *gan* “particularly” occurs with certain verbs, grouping them into the following categories, although not reaching any conclusions about the relationship between the above types of verbs and the meaning of *gan*: a) motion; b) vision; c) sound and noise; d) desire and entreaty; and e) emotion. Likewise, but in relation to *can* and *couth* in Scots, Taylor (1917: 576) lists the results of his vocabulary test, but he seems to be more interested in how frequently certain verbs are with both variants in this variety (provided in (19) below for reference). Finally, Kerkhof (1966: 156) also provides the same categories of verbs with *gan* in ME as Mustanoja (1960), but makes an observation that *gan* is found with “some verbs denoting the perfective and resultative aspect”, probably referring here to verbs denoting a terminal point, namely achievement verbs.

(19)	Barbour	<i>Tak</i> 30; <i>ga</i> 25; <i>mak</i> 21; <i>say</i> 12; <i>cry</i> 11; <i>sla</i> 7; <i>say</i> 6; <i>se</i> 6.
	Wyntoun	<i>Pas</i> 8; <i>mak</i> 9; <i>ga</i> 5; <i>say</i> 9; <i>tak</i> 3; <i>se</i> 3.
	Blind Harry	<i>Far</i> 16; <i>gang</i> 12; <i>tak</i> 18; <i>pas</i> 14; <i>ride</i> 9; <i>draw</i> 14.
	Holland	Data too slight to be of any value.
	Henryson	<i>Tak</i> 6; <i>cry</i> 6; <i>declair</i> 3; <i>pas</i> 3; <i>hyng</i> 3; <i>call</i> 3; <i>sing</i> 3; <i>say</i> 3; <i>crepe</i> 3. (Taylor 1917: 585)

While the above accounts make some observations on the categories of verbs found with *gan* and *can* in ME and Scots, the categories identified by such sources do not reflect distinctions between properties of different types of events conveyed by verbal expressions. For this reason, other classifications of verbs in the infinitive found with *gan* and *can* have been applied, ranging from two-way (perfective vs. non-perfective, punctual vs. non-punctual) to four-way (‘states’, ‘activities’, ‘accomplishments’ and ‘achievements’) (cf. Vendler 1957). For example, Beschorner (1920: 13-14) effectively discusses two groups of verbs resulting in the non-ingressive meaning of *gan* and *can* in ME: verbs denoting instantaneous or temporary action, as in (20) below, and verbs expressing a continuing action, as in (21) below, with these groups corresponding to punctual and durative, respectively. Next is Funke (1922: 12, 24; translation mine), who sees “Perfektiv” [perfective]

verbs in infinitive (a perfective verb is seen by Funke 1922 as punctual, i.e. not analysable into a beginning, middle and end) as contributing to a non-ingressive meaning of *gan* (i.e. “Deskriptiv” [descriptive] in his terminology), but only sporadically pointing this out, for example, by saying that “*ariue* ist doch perfektiv” [*arrive* is perfective], as in (22) below. Similar proposals to the ones listed here are also made by Wuth (1915: 55) and Kerkhof (1966: 154-158), although using different terminology, whereas those by Brinton (1981; 1983; 1988; 1996) and Sims (2008) apply the four-way classification of verbs to their findings.

- (20) *And how his ancestre. African so dere,*
And how his ancestor. African so dear,
Gan in his slepe that night to him appere
Did in his sleep that night to him appear
 (Chaucer *The Parliament of Fowls* 42; Beschorner 1920: 13)

‘And how his ancestor, African so noble, appeared to him in his sleep that night’

- (21) *And from the morwe gan this speche leste*
And from the morn did this speech last
Til downward drow the sonne wonder faste
Till downwards drew the sun wondrously fast
 (Chaucer *The Parliament of Fowls* 489; Beschorner 1920: 14)

‘And this speech lasted from the morning until the sun went down’

- (22) *A king þer gan ariue*
A king there did arrive
Pat wolde hire haue to wyve⁹⁰ ...
That would her have to marry ...
 (*King Horn* 921; Funke 1922: 14)

‘A king arrived there, wishing to marry her’

Using findings by Wuth (1915: 55), Beschorner (1920: 13-14), Funke (1922: 12, 24) and Kerkhof (1966: 156-157), Brinton (1981: 169-174; 1983: 239-241; 1988: 154-1616; 1996: 75-78) offers diagnostics to distinguish between non-ingressive and ingressive meanings of *gan* in ME. However, these diagnostics differ from those

⁹⁰ In Funke (1922: 14), this word occurs as <wyne>, whereas in the *TEAMS* as <wyve> and in the *Corpus of Middle English Prose and Verse* as <wyue>. It, therefore, appears that <wyne> is a typographical mistake.

proposed by Wuth (1915), Beschorner (1920), Funke (1922) and Kerkhof (1966) because they are based on the effects of linguistic context on an aspectual interpretation of various verbs in ME texts. Brinton (1983: 238), for example, argues that in ME, the verb *gan* occurs with a “non-aspectual” reading, “[i]f the infinitive is a perfective verb ... or an achievement, then it expresses a punctual event, one which happens in a moment and which is not temporarily analyzable into a beginning, middle, or end” (cf. Brinton 1981: 170). The meaning of this verb is non-ingressive in such cases, because “[i]t is impossible to focus on the inception of such an event, since the moments of inception and of termination coincide.” Brinton (1981: 170) maintains that only two instances of verbs quoted by Beschorner (1920) are “truly punctual”, namely, “(*gan* she *fine*, *gan* ... *misse*)”, resulting in the non-ingressive sense of *gan*. She also includes other examples of such verbs from sources like Funke (1922), Koziol (1932), Kerkhof (1966) and Mustanoja (1960): *meten* (*imeten*) ‘to meet’, as (23) below, *fongen* ‘to seize’, *espyen* ‘to catch sight of’, *missen* ‘to miss’, *bresten* ‘to burst’, *ariven* ‘to arrive’, *kennen* ‘to recognise’, *nimen* ‘to seize’, *senden* ‘to send’, *quellen* ‘to kill’, and *ap(p)eren* ‘to appear’.

(23) *Horn* ... *ferde* *to* *wude* *for* *to* *schete*;
 Horn ... went *to* wood *for* to shoot;
 A *knave* *he* *gan* *imete*
 A knave *he* did meet
 (*King Horn* 940; Mustanoja 1960: 611, as cited in Brinton 1981: 171)

‘Horn went to the woods to shoot; he met a peasant’⁹¹

However, not all verbs found with *gan* and *can* in ME and Scots are punctual verbs. Brinton (1981: 170), for example, argues that the verbs that Beschorner (1920: 14) cites, i.e. *calle*, *finde*, *falle*, *take*, *dye* and *misse*, permit an ingressive interpretation of *gan* and *can* because *falle* is a process verb, *dye* and *take* development verbs, and *calle* an iterative verb. Given the fact that *gan* and *can* are allegedly ingressive in such cases, this verb and its variant are similar in meaning to verbs like *beginnen*, which normally denote ingressive aspect, and are compatible with verbs expressing a durative event, i.e. an activity or accomplishment, or an iterative series, i.e. a state

⁹¹ Brinton (1981: 171) translates the above as ‘Horn went to the woods to shoot; a knave he met (NOT began to meet)’; I have provided an alternative translation.

(Brinton 1983: 238-239; 1988: 119-120; 1996: 67-69). Furthermore, there are contexts when a punctual verb in infinitive found with *gan* can result in an ingressive meaning. For example, Brinton (1988: 157) argues that this kind of situation occurs “[w]hen the verb names a punctual situation which can be repeated and when the object or subject of [the punctual verb] is a plural count noun or a mass noun”. She maintains that “[t]he repeated punctual events constitute a series, which is durative; one can then focus on the beginning of the series”. Brinton (1981: 172) provides instances of punctual verbs in infinitive, such as *quelle* ‘to kill’, as in (24) below, *felle* ‘to fell’, *mete*, *aryue*, *deffie* ‘to defy’, *sle* ‘to slay’ found with *gan*, where subjects and objects are said to affect the meaning of this verb, resulting in an ingressive interpretation. As the subject in (24) below is in the plural, the event or situation of *quelle*[ing] can be viewed here as a series, i.e. seen performed by each person individually. In this example, therefore, one can perhaps deal with the verb *quelle* as a semelfactive.

- (24) *Pat folc hi gunne quelle*
 That folk they began to kill
And churchen for to felle
 And churches for to fell
 (*King Horn* 61-62; Funke 1922: 13, as cited in Brinton 1981: 172)

‘They began to kill people and demolish the churches’⁹²

Other linguistic context can also have an effect on the meaning of *gan* and *can* in ME and Scots. For example, helpful in isolating non-ingressive instances of *gan* are especially adverbials of duration or iteration (Brinton 1981; 1988; 1983; 1996). Brinton (1988: 157) argues, “adverbs expressing duration and those expressing iteration are opposed to ingressive aspect, which makes reference to the single moment of beginning.” Only adverbials linked to time-at-which are compatible with ingressive aspect (Brinton 1981: 173; 1983: 240; 1988: 157-158, cf. Dowty 1979, C. S. Smith 1991, Piñón 1997). Following from this, if found with adverbials of duration or iteration, *gan* and *can* are likely to have a non-ingressive meaning. In contrast, if occurring with adverbials of time-at-which, *beginnen* is likely to be used,

⁹² Brinton (1981: 172) translates the above as ‘They began to kill people and to fell the churches’; I have provided an alternative translation.

given that it conveys ingressive meaning. Brinton (1981: 173) then collates examples of *gan* and *can* with durative and iterative adverbials from such sources as the *MED*, Beschorner (1920), Funke (1922), Koziol (1932) and Mustanoja (1960), adding that “[p]articularly striking are [those] which are antithetical to a focus on the moment of inception”, in which case they would help ascertain the non-ingressive meaning of this verb and its variant in ME: *ful oft* ‘very often’, as in (25) below, *mony dayes* ‘many days’, *a fourtenizt* ‘a fortnight’, *ay* ‘always’, *day by day* ‘day by day’, *fiftene ere* ‘fifteen years’, *fyue wynter* ‘five winters’, and *bope nizt & dai* ‘both night and day’.

- (25) *his cher ful oft con chaunge/*
 his mood full often did change/
at chapel er he myzt sene
 that chapel before he might see
 (*Sir Gawain and the Green Knight* 711-712; Mustanoja 1960: 611, as cited in Brinton 1981: 173)

‘his mood changed very often before he saw that chapel’,⁹³

Brinton’s (1981; 1983; 1988) diagnostics may be helpful in determining whether *gan* and *can* are used with the non-ingressive meaning, this, therefore, being applicable to this verb and its variant in the NME and ESc data sets. The proposal made by this source seems to acknowledge the fact that an aspectual interpretation of a verb takes into account the linguistic context, rather than just the basic ‘Aktionsart’ of the verb in infinitive. However, there may be difficulty in interpreting some verbs as punctual. In particular, some verbs may be more punctual than others, but also it is possible for them to change categories, depending on the context, as exemplified earlier in this section to be the case for the verb *run* in ModE. Furthermore, Brinton (1981: 172; 1983; 1988) argues that “it [just] seems easier” to interpret the meaning of *gan* as ingressive if the subject or object of this verb with a punctual (achievement) verb in infinitive is a plural count noun or a mass noun, with this diagnostic not mentioned in Brinton’s (1996) more recent account on the grammaticalization of this verb in ME. In such cases, I would maintain that punctual (achievement) verbs could also be

⁹³ Brinton (1988: 173) translates the above as ‘his face changed (NOT began to change) very often before he might see the chapel’; I have provided an alternative translation.

effectively reinterpreted as semelfactive verbs, with the event or situation denoted by them as constituting a series, characterised by duration. However, I also show in this chapter that only in some cases, it may be helpful to establish the non-ingressive meaning of *gan* and *can* in the NME and ESc data sets on the basis of this diagnostic.

So far, I have concentrated on reviewing the literature dealing with the meaning of *gan* and *can* in ME. However, Brinton (1981) identifies instances of ME *beginnen*, while Los (2000; 2005) examples of OE *beginnan*, where the meaning of this verb can be interpreted as non-ingressive. Both sources attempt to account for this kind of meaning, although each differently.⁹⁴ Brinton (1981: 184-185) relies on the *MED*'s, Häusermann's (1930) and Visser's (1969) data in analysing instances of ME *beginnen* in the non-ingressive sense, giving an interpretation of the meaning of this verb as based on "the context as a whole" (Brinton 1981: 185). She interprets *bygynneþ* in (26) below as non-ingressive, whereas that of *gan* in the same example as either ingressive or non-ingressive; she claims, "[t]he first *biginnen* periphrase must, of course, have a generic interpretation, but the second might allow an ingressive meaning." However, I would argue that Brinton's (1981) perception of *bygynneþ* here as being non-ingressive is only probably due to the fact that *sprynge* could refer to the act of jumping, especially when considered in isolation and in combination with *beginnen* being in the present tense, rather than a number of individual jumps, in which case it would be a semelfactive verb. Therefore, I would maintain that an interpretation of whether the meaning of *beginnen* is ingressive or non-ingressive would be based on the treatment of the verb in infinitive as either punctual (achievement) verb or semelfactive verb – the action denoted by the verb *sprynge* seems to be taking place in the background and so is initiated, but not complete, before the action or event of *he gan aboute the chaumbre sterte* happens. Moreover, the presence of adverbials of time and place *Now her, now ther*, appears to suggest that the action or event denoted by the verb in infinitive is not complete, but on-going. This suggests that the verb *bygynneþ* would, in fact, convey the idea

⁹⁴ Los' (2000; 2005) proposals in relation to the choice of infinitive and semantic bleaching have already been reviewed in Chapter Four. They will not, therefore, be repeated here.

of ingressiveness here, with the present indicative form sometimes being used instead of the preterite one in descriptions (Richardson 1991a).

(26) *Right as the wylde bole bygynneth sprynge,*
 Right as the wild bull does spring,
 Now her, now ther ...
 Now here, now there ...
 Right so gan he aboute the chaumbre
 Right so did/began he about the chamber
 sterte
 leap

(Chaucer *Troilus and Criseyde* 4.239; Brinton 1981: 185; translation in the original)

‘just as the wild bull springs (NOT begins to jump), now here, now there ... so he began to leap (or he leapt [sic]) around the chamber’

In sum, although there are some limitations to proposals made in the literature on the issue of identifying the meaning of *gan* and *can* in ME and Scots, distinguishing punctual (achievement) verbs from other types of verbs by taking into consideration their Aktionsart properties can be a diagnostic in helping to determine the meaning of *gan* and *can* in ME and Scots. Furthermore, using adverbials of duration or iteration as a diagnostic may also help when attempting to determine the meaning of the above verb and its variant, with an outcome that when *gan* and *can* are found with such an adverb, they are likely have a non-ingressive interpretation. However, when *gan* and *can* with a punctual (achievement) verb in infinitive are used with a plural count noun or mass noun, an ingressive meaning of this verb and its variant is likely, although this diagnostic is omitted in Brinton (1996). Given the universality of the proposals outlined in this section, I apply to my analysis Brinton’s (1981; 1983; 1988; 1996) diagnostics, with the expectation that if *gan* and *can* are used with punctual (achievement) verbs or with adverbials of duration or iteration, the meaning of this verb and its variant will be non-ingressive in the NME and ESc data sets. However, I also advocate the use of other contextual information, i.e. clues provided in a description by the author of a text – as a useful way of helping to work out whether this verb and its variant are used with non-ingressive or ingressive meaning in the NME and ESc data sets. Finally, I discuss the change in meaning from ingressive to non-ingressive in terms of mechanisms and processes of

grammaticalization. I, therefore, show that *gan* and *can* in these data sets are a development from OE *onginnan* and OE and eME *aginnan*, originally expressing ingression, but shown in the literature to have undergone semantic bleaching in OE and in eME period when used with the plain infinitive.

5.4 Changes in the Meaning of *Gan*, *Can* and *Beginnen* in Northern Middle English and Early Scots

The notion of meaning change goes back to Paul ([1880] 2009), who proposes a classification of meaning change through generalisation, metaphor, metonymy and specification. Others, such as Stern (1931), Ullmann (1963) and A. Blank (1997), provide more elaborate theories on semantic change. In addition to these, Lakoff & Johnson (1980) and Indurkha (1992) offer investigations into this area by metaphor; Panther & Radden (1999) by metonymy; and Dik (1977) and Geeraerts (1997) by prototype shifts. While it is generally accepted (cf. de Saussure 1962) that meaning change is driven by language external factors, it has also been proposed that there are areas of meaning change which are motivated by linguistic factors. For example, Hopper & Traugott ([1993] 2003) and Levinson (2000) argue that semantic change is motivated by pragmatic inferencing and then semanticization, while Eckardt (2006) looks at meaning change under reanalysis in the paradigm of truth conditional semantics. In what follows, I discuss meaning change associated with the process of grammaticalization, namely semantic bleaching.

5.4.1 Semantic Bleaching

Meaning change is particularly evident in grammaticalization, although it is not restricted to this process. As grammaticalization continues, semantic bleaching occurs, with grammaticalized forms losing pragmatic and semantic meaning, having become more syntacticized and morphologized. This loss of meaning is the result of reanalysis, or as Harris & L. Campbell (1995: 92) rather put it, “it is the essence of the reanalysis itself”. It occurs nearer the end of the process of grammaticalization, i.e. after reanalysis⁹⁵ (cf. Sweetser 1988; 1990, Haiman 1991, Traugott & König 1991, Hopper & Traugott [1993] 2003), although other sources have claimed that a

⁹⁵ However, this also depends on the type of development, with any kind of change being regarded as nearer the ‘end’, when the idea of the cline of grammaticality is taken into account.

loss of meaning is the cause of, or pre-requisite to, grammaticalization, too (cf. Bybee, Perkins & Pagliuca 1994, Rubba 1994) (although Fischer 1994, using examples of the ModE verb *have to* and OE and ME changes of word order, shows that semantic bleaching does not direct this process). Furthermore, the loss of semantic content does not happen suddenly, as the process of grammaticalization is characterised by gradualness. This involves a number of discrete micro changes (Hopper & Traugott [1993] 2003: 77, cf. Traugott 2012a; 2012b) or “tiny little steps” (Brinton & Traugott 2005: 150), in some cases leading to “abductive change” (Lightfoot 1979). Next, details of the lexical history of the item undergoing grammaticalization may be shown in restrictions on its grammatical distribution (Hopper & Traugott [1993] 2003: 96). In other words, traces of the original meaning of a form undergoing grammaticalization from lexical to grammatical can still be found (cf. Hopper 1991), but there is also the promotion and demotion of some lexical meanings (those that are promoted have a tendency to be abstract) (Hopper & Traugott [1993] 2003: 71-98).

Grammaticalization involves loss of semantic content and without doubt, meanings tend to weaken over time during this process. The initial stages of semantic bleaching are characterised by redistribution or shift of meaning, rather than a loss of meaning (Hopper & Traugott [1993] 2003: 94). Brinton (1988), for example, argues that changes in meaning are central to the process of grammaticalization and involve a metonymic shift in focus. She, then, focusses on what she refers to as the semantic aspect of grammaticalization, and so probably semantic bleaching, when discussing the development of aspectualizers (ingressive, egressive and continuative) in English, saying that “grammaticalization should be considered a semantic rather than syntactic process” (Brinton 1988: 95). Brinton (1988: 111) also refers to auxiliation, i.e. change from a lexical verb to an aspectualizer, associating this process primarily with changes in meaning, too “[v]iewed as a semantic process, auxiliation raises two issues: one about the choice of verbs to become auxiliaries and the other about the semantic changes occurring in these verbs when they are auxiliated”. Therefore, Brinton (1988) claims that grammaticalization concerns changes in the meaning of expressions of aspect as central to the process, but while this may be the case, it is

also important to remember that meaning change may occur without grammaticalization. Brinton (1988: 95) argues,

[t]he choice of aspectualizers in the history of English can be explained by the observable conformity between the spatial meanings of aspect categories and the semantics of the verbs involved: going into a situation (ingressive aspect) is marked by verbs expressing movement into; being in a situation (continuative/iterative aspect) is marked by verbs expressing location in; and coming out of a situation (egressive aspect) is marked by verbs expressing movement out of. Moreover, since aspect categories are spatial, the semantic change affecting a verb during grammaticalization is a metonymic shift from one spatial meaning to another, not a gradual bleaching from spatial to aspect meaning.

When considering the meaning of *gan* and *can* on the one hand and that of the etymologically related *beginnen* on the other, the former two appear to show characteristics of semantic bleaching in ME and Scots in that they are no longer associated with the idea of ingressiveness, but are found with a non-ingressive meaning instead (Wuth 1915, Beschorner 1920: 13-14, Funke 1922: 14, Mustanoja 1960: 613, Kerkhof 1966: 156-158, Visser 1969: 1573-1574, 1575-1577, Brinton 1981: 161, 170-173; 1983: 236; 1996: 67-83, Denison 1993: 322, Sims 2008: 171). For example, Funke (1922: 15; translation mine) uses the expression “farbloses Flickwort [colourless filler]” to refer to semantically bleached instances of *gan*, whereas Mustanoja (1960: 611) states, “[t]hrough frequent use [*gan* with infinitive] loses its ingressive colour more and more”. In a similar way, Visser (1969: 1571) suggests that the loss of meaning of *gan* is gradual and on-going in ME, “*gan* ... in combination with an infinitive came to be used in a weakened sense, lost its meaning of ‘commence’ and gradually approached or passed into an empty, periphrastic verb, semantically equivalent to periphrastic *did*”. Brinton (1983: 235-236) explicitly states that “[i]n many cases [of *gan* with infinitive] ... the ingressive meaning is bleached”, whereas in Brinton (1996), she links the (meaningful) textual function and the (meaningless) metrical function to the complete loss of meaning of *gan*, “in its later development ... *gan* seems to have been coopted as a metrical expedient ... [t]his move ultimately lead[ing] to *gan*’s complete loss of meaning and disappearance, a movement to \emptyset marker compatible with the known path of grammaticalization” (81-82). Overall, therefore, the meaning of *gan* has frequently been described in the literature as semantically bleached.

However, it is already possible to observe semantically bleached instances of OE *onginnan* and OE and eME *aginnan*. This non-aspectual meaning is found when both verbs occur with punctual (achievement) verbs in infinitive, as in (27) below and (28) below (Brinton 1981: 169-170; 1983: 238; 1988: 159-160, cf. Wuth 1915, Funke 1922, Mustanoja 1960, Visser 1969: 1580) (cf. Los 2000; 2005 on the grammaticalization of OE *onginnan* and OE *beginnan* with the plain infinitive in Ælfric). OE *onginnan* is also non-ingressive when it is found with “an iterative adverbial genitive” and/or “a durative adverbial clause”, as in (29) below (Brinton 1988: 160-161, cf. Brinton 1981: 173; 1983: 240).⁹⁶ Despite the above-mentioned changes in relation to the meaning of these verbs, Brinton (1988: 158, 159-161) and Sims (2008: 121) point out that the non-ingressive use of *onginnan* and *aginnan* is minimal in OE, with Brinton (1988: 117) adding that the use of these verbs as markers of ingressive aspect is still common at the time. However, Los (2000: 256; 2005) finds that OE *onginnan* was more frequently found with the plain infinitive in Ælfric (18 instances), than with the *to*-infinitive (seven instances), especially that “[t]he bleached, ‘pleonastic’ instances of these ingressive aspectualizers only occur with the bare infinitive” (Los 2000: 259). Regardless of how frequent these verbs are, it is already possible to interpret instances of OE *onginnan* and OE and eME *aginnan* in (27) below, (28) below and (29) below as portraying an action or event as occurring in the past. Furthermore, such instances of these verbs also depict the action or event as being complete, i.e. whole, instead of focussing on the inception of, or entering into, an event or situation denoted in the verb in infinitive, where the aspectual interpretation of such cases would be understood as denoting an action or event as either complete or incomplete. This interpretation can also be applied to non-ingressive instances of *gan* and *can* in the NME and ESc data sets, as will be discussed below.

(27) *ge* *þa* *sciran miht* *deman* *ongunnon*
 you then bright might judge did

⁹⁶ Brinton (1981: 173-174; 1983: 240-241; 1988: 158-161) lists no examples of OE and eME *aginnan* with adverbials of duration or iteration. However, the lack of instances of this variant with such adverbials in OE and eME could be due to accidental gap, given that instances of *onginnan* and *gan* are found in OE and ME, respectively, with adverbials of duration or iteration.

(*El* 310b-311a; Brinton 1988: 160; translation in the original)

‘you then judged (NOT began to judge) the bright power (of Christ)’

- (28) *Ʒa þe cniht wende þat hit þe eotend*
Then the knight weened that it the giant
weorenand he an-bursten agon swulc weore a wilde
were and he enrage did so were a wild
bar
boar
(Lazamon *Brut* 25831-25834; Brinton 1988: 158)

‘Then the knight thought that it was the giant and he enraged like a wild boar’,⁹⁷

- (29) *Ongan þa dryhtnes æ dæges ond nihtes ...*
Did then God’s law days and nights ...
georne cyðan
eagerly exhibit
(*El* 198a-9b; Brinton 1988: 161; translation in the original)

‘[He] then eagerly made known (NOT began to make known) the law of God day and night’

In this section, I discuss non-ingressive instances of *gan* and *can* found in the NME and ESc data sets. Therefore, like with OE *onginnan* and OE and eME *aginnan*, the non-ingressive meaning of *gan* and *can* in these data sets is mostly brought on when they occur with punctual (achievement) verbs in infinitive, as in (30) below, (31) below and (32) below (only a selection of examples will be discussed in the remainder of this chapter, due to the spatial limits of this study). For example, the verb *reu* found with *gan* in the NME data set refers to a punctual event or situation meaning ‘to repent of (sins or of crimes freq. regarded theologically); to feel penitence, remorse, or contrition for; to do penance for’ (*OED* s.v. *reu* v. 3c). In addition, contextual information also supports this punctual reading: events or situations in this story are sequential and conditional in that in the preceding context, not listed in this example, Jesus tells the Samaritan woman that only through him, one can obtain and receive eternal life, i.e. salvation. This salvation comes only to those who confess,

⁹⁷ Brinton (1988: 158) translates the above as ‘then the knight went to the place where the giants were and he burst out (NOT began to burst out) just as if he were a wild bear’. However, the noun *bar* refers to ‘an uncastrated male swine (either wild or domesticated)’ (*MED* s.v. *bōr* n. 1), with the expression *brust (iburst) as a bor* meaning ‘bristly as a boar, bristling (or showing anger) like a boar’ (*MED* s.v. *bōr* n. 1d), rather than a bear. I have provided an alternative translation.

and repent, their sins and so desire forgiveness. Before the Samaritan woman can embrace the love of Jesus, she has to repent not some, but all sins. Renouncing all sins is presented to the reader as referring to its entirety, especially by the use of the personal pronoun *hir* ‘her’ and the plural *dedes* ‘deeds’; it denotes the idea of completeness. Overall, therefore, it seems that primarily because of the Aktionsart of the verb in infinitive, but also because of contextual information, **gan** and **can** have a non-ingressive meaning (designating a complete event or situation denoted by the verb in infinitive) in the examples referred to in this paragraph.

- (30) *Þe woman þan hir dedes gan rew*
 The woman then her deeds did repent
And þus scho said vnto Jesu,
 And thus she said unto Jesus,
 (1350-1420 *The Northern Homily Cycle* P II p. 82; *HCET*)

‘The woman then repented her sins and said to Jesus in this way,’

- (31) *Wit þat stan he laid in sling,*
 With that stone he laid in sling,
Sua stalworthli he lete it suing
 So strongly he let it swing
Pat in his frunt þat stan he fest,
 That in his forehead that stone he cast,
Pat bath his eien vte can brest
 That both his eyes out did burst
 (1350-1420 *Cursor Mundi* p. 438; *HCET*)

‘He put that stone in the sling, and let it swing so strongly that he cast it on the forehead, that both of his eyes came out’

- (32) *A new dysese thane can he tak,*
 A new hardship then did he suffer,
Seand his wyf sic sorou mak;
 Seeing his wife such sorrow make;
 (1375-1420 *The Life of St. Julian Hospitaller* l. 386-390; *TEAMS*)

‘He then suffered new hardship, seeing that his wife brought so much sorrow;’

Duration and iteration adverbials can have an effect on the meaning of **gan** in the NME data set, too (cf. Brinton 1981: 173; 1983: 240; 1988: 157). For example, the non-ingressive meaning of **gan** in (33) below seems to be brought on by *are*, in this

context meaning ‘during some former time; formerly, previously, before’ (*MED* s.v. *ērer* adv. (comp.)). This adverbial appears to specify duration in the past when the event or situation denoted by the activity verb *vse* ‘use’ in infinitive took place, resulting in *gan* denoting the completeness of this event or situation in its entirety, rather than signifying that the event or situation is simply finished (the presence of the determiner *All* also supports the idea of perfectivity or completeness here – it denotes the idea of the whole of something). By comparison, however, the presence of the adverbial of time-at-which *Pe morn* ‘The next morning’ in (34) below (the punctual (achievement) verb *dight* ‘saddle’ on its own appears to be compatible with adverbials of time-at which, cf. Dowty 1979: 58-59) does not appear to affect the meaning of *can*. Here, it seems that the non-ingressive meaning of this variant can be worked out on the basis of the use of the punctual (achievement) verb *dight* and other contextual information, i.e. a description of the sequence of events, indicating that once the event or situation of saddling the donkey was complete, the protagonist headed for Jerusalem. Overall, therefore, although the Aktionsart of the verb in infinitive can help determine that *gan* and *can* in the NME and ESc data sets are used with non-ingressive meaning, the presence of adverbials of duration or iteration and contextual information appears to play a significant role in establishing the meaning of this verb and its variant, too.

- (33) *All vices þat scho are gan vse,*
 All vices that she before did use,
All turned scho þan into vertuse
 All turned she then into virtue/morality/strength
 (1350-1420 *The Northern Homily Cycle* P III p. 127; *HCET*)

‘All the vices that she possessed during that time, she turned all into virtue’

- (34) *Ful wel war þai wit him þat night*
 Full well were they with him that night
Pe morn his ass þan can he dight,
 The morn his ass then did he saddle,
To ierusalem he tok þe strete
 To Jerusalem he took the road
 (1350-1420 *Cursor Mundi* pp. 1011-1012; *HCET*)

‘They were able to be with him that night. The next morning he saddled his donkey, and took the road to Jerusalem’

Finally, contextual information on its own (rather than the verb in infinitive, adverbials of duration or iteration) can help establish that the meaning of some instances of *gan* and *can* in the NME data set is non-ingressive. For example, *can* is used with the verb *vnderstand* ‘understand’, as in (35) below, which in itself can be classed as a temporary state. However, the event or situation described in the subsequent line, namely *To laban tald sco new tipand* ‘To Laban told she new tiding’, implies that the character here must have understood or accepted that Jacob would be her husband before telling Laban about it. Therefore, the grasping of the news by the woman has aspectual restrictions imposed on it; in that it must be understood to be complete before the event or situation of telling the news takes place. Furthermore, the example of the verb *vnderstand* with the variant *can* was also listed in the *OED* (s.v. *can* v.² 2a) before the entry was upgraded, where it occurred under the sense ‘[i]t was usually followed by an infinitive without *to*, and then approached or passed into a simple auxiliary of the past tense = the modern *did*’ (the updated entry of *can* as a variant of *gan* contains the following: ‘[f]ollowed by bare infinitive, as a periphrastic auxiliary of the past tense’: ‘[w]ith infinitive of a main verb, e.g. *tho can she weepe* ‘then she wept’; = *did*’ *OED* s.v. *can* v.² I 1a). However, Ogura (1998: 301) provides the same example, but translates it as either ingressive or non-ingressive, as in (36) below, although she does not provide the reasons for this. More importantly, though, the difference in the use of *can*, *gan* and *bigon* in Ogura’s (1998) example lies in whether the event or situation described here is presented as complete or not.

- (35) “*I am iacob, þi cosin nere,*
 “I am Jacob, your cousin near,
For þi luue am i comen here.”
 For your love am I come here.”
þai mai quen sco can vnderstand
 the womanwhen she did understand
þat iacob suld be hir husband
 that Jacob should be her husband
To laban tald sco new tipand,
 To Laban told she new tiding,
 (1350-1420 *Cursor Mundi* p. 227; *HCET*)

“I am Jacob, your cousin, for your love I have come here.” When the woman understood that Jacob should be her husband, she told Laban the new news,’

- (36) C: *þai mai queen sco can vnderstand*
G: *þe may, quen scho gan vnderstand*
T: *Whenne she bigon to vndirstonde*

‘the (who) can, when she understood’ or ‘When she began to understand’

(*Cursor Mundi* 385; Ogura 1998: 301; translation in the original)⁹⁸

There are other instances of *gan* and *can* where contextual information may provide vital clues that this verb and its variant are used in the non-ingressive sense in the NME and ESc data sets. For example, *gan* is found with the verb *tell*, as in (37) below, which without context can be seen as either non-punctual or punctual. However, in this instance, the use of the NP *all þe tale* ‘all the tale’ provides further information in that the event or situation of telling the tale is conveyed to the reader as complete, since it involves the determiner *all* and a singular count noun. However, although Brinton (1981: 172; 1983: 240; 1988: 157) proposes that in certain contexts, an ingressive interpretation is likely when the subject or object of *gan* and *can* with a punctual (achievement) verb in infinitive is a plural count noun and a mass noun, this is not the case for *can* in (38) below, where the plural NP *dedis* ‘actions’, functioning as the object of the activity verb *dele* ‘arrange’, refers to God’s plans that have already been devised, as becomes apparent from the subsequent lines: *þat ioseph suld be funden þus þat prisund was, and noght iesus* ‘that Joseph should be found, thus, [who] was imprisoned, and not Jesus’. That the plans have already been devised is also suggested by the NP *þis tþand* ‘this tiding’, especially the determiner *þis*, pointing to something specific and definite, and the VP *thanked þair godd of israel* ‘thanked their God of Israel’, referring here to the act of expressing gratitude to God for something that has already been carried out, rather than something that will be carried out or is being carried out. Therefore, *gan* and *can* in these examples appear to be used with the non-ingressive meaning, with the actions or events denoted in the

⁹⁸ The NP *þai mai* in the Cotton Vespasian A.3 MS and *þe may* in the Göttingen MS Theol. 107 refer to ‘a woman; esp. a young attractive woman (married or single)’ (*MED* s.v. *mai* n.¹), rather than “the (who) can”, as shown in Ogura (1998).

verb in infinitive intended to be conveyed as complete, this meaning established on the basis of contextual information in the text.

- (37) *When þe monk wist how he had bene
 When the monk was aware how he had been
 Fra hame thre hundreth zere bedene,
 From home three hundredth year altogether,
 Pan all þe tale he gan þam tell
 Then all the tale he did them tell
 What made him so lang to dwell
 What made him so long to dwell
 (1350-1420 *The Northern Homily Cycle* P II p. 205; *HCET*)*

‘When the monk became aware how he had been [away] from home for three hundred years altogether, he then told them the story about how he came to live so long’

- (38) *All on þis tīþand wonder þai,
 All on this tiding wonder they,
 And thanked þair godd of israel,
 And thanked their God of Israel,
 Pat sua-gat can his dedis dele
 That sogate did his actions arrange
 (1350-1420 *Cursor Mundi* p. 1009-1010; *HCET*)*

‘They wondered about this news, and thanked their God of Israel for the way he arranged his plans’

Other examples where contextual information plays a significant role in determining the meaning of *gan* and *can* in the NME and ESc data sets may involve state verbs, following Vendler’s (1957; 1967) classification of verbs. (State verbs can be interpreted as either permanent or temporary states. Permanent states are different from temporary states in that the former are atelic, whereas the latter are telic, although both still remain durative.) The idea of a permanent state entails the temporary durativity of this state over any given point of reference, though the idea of a temporary durative state at a point of reference conversationally suggests that the state is temporary. For example, it appears that the state verb *dwell* in (39) below refers to a particular moment or period in time when the NP *many leprus men* ‘many impure men’ were already in *þe landes of Israell* ‘the lands of Israel’. Contextual information supports this temporary interpretation of the state verb here: although

there were many widows in the Israel in the time of Elisyus, Elys was not sent to any of them, but only to a poor woman in Sydoyne; at the time, many lepers were already in Israel, but none of them were cleansed, except for Naman Syrus, who was a Syrian. In this case, therefore, *gan* is used in the non-ingressive sense, with the contextual information helping to establish that the verb *dwel* found with it refers to a punctual event or situation. A similar case arises with the verb *won*, as in (40) below, where the presence of the clause *And gaf it sepin to Josep his sun* ‘And gave it then to Joseph, his son’ implies that the period during which Jacob lived in *þe felde* ‘the field’ beside *a cete þat hight Sichar* ‘a city called Sichar’, mentioned earlier in the text, but not quoted here, should be understood as finished, the event or situation presented here being complete. Overall, the role of the contextual information is crucial and vital in determining the meaning of some instances of *gan* and *can* in the NME and ESc data sets.

- (39) *So þe buke of kinges beris testimoin.*
 So the book of kings bears testimony.
And many leprus men gan dwel
 And many impure men did dwell
þan in þe landes of Israell
 Then in the lands of Israel
þat lifed þan and war leprus
 That lived then and were impure
Vnder þe prophet Elisyus,
 Under the prophet Elisyus,
 (1350-1420 *The Northern Homily Cycle* P II p. 68; *HCET*)

‘So the book of kings bears testimony. And many impure men dwelt in the lands of Israel; that lived then and were impure under the prophet Elisyus,’

- (40) *Jacob it boght of Emar suns,*
 Jacob it bought of Emar sons,
And þarin sum tyme gan he won
 And therein some time did he live
And gaf it sepin to Josep his sun
 And gave it then to Josep his son
 (1350-1420 *The Northern Homily Cycle* P II p. 77; *HCET*)

‘Jacob bought [the plot of ground] from sons of Emar, and he lived there some time and gave it then to Joseph, his son’

Overall, therefore, the Aktionsart of the verb in infinitive and adverbials of duration or iteration, but also contextual information in a text, can provide clues about the meaning of *gan* and *can* in the NME and ESc data sets.

5.4.2 Opacity

As already referred to throughout this chapter, some sources are in disagreement as to whether particular instances of *gan* and *can* in ME and Scots are semantically bleached or lexically meaningful (cf. Eienkel 1891, Taylor 1917: 55, Kerkhof 1966: 158-159, Visser 1969: 1379, Brinton 1981: 161, Denison 1993: 322). This kind of view is also reflected in the findings in the NME and ESc data sets, where some instances of *gan* and *can* are categorised as ambiguous, given that they permit either an ingressive or non-ingressive interpretation, a situation known as opacity. Opacity is recognised to occur in the process of grammaticalization, where in some contexts, two interpretations or analyses are possible (cf. Hopper & Traugott [1993] 2003: 52, Diewald 2002: 116, Diewald & Smirnova 2012: 113-114, 116-117, 125 on semantic opacity and its examples). However, the presence of semantic opacity could also be seen as obscuring the establishment of *gan* and *can* as auxiliary constructions. Opacity is also different from persistence in that it allows “for the structure to continue to be analysed as before, and for a new analysis to be innovated, and then to coexist with the earlier analysis [i.e. both analyses continuing to exist, but with different meanings]” (Hopper & Traugott [1993] 2003: 52), whereas in the case of persistence, “when a form undergoes grammaticalization from a lexical to a grammatical item, some traces of its original meanings tend to adhere to it, and details of its lexical history may be reflected in constraints on its grammatical distribution”. Overall, therefore, some instances of *gan* and *can* in ME and Scots do not neatly fit into the categories or descriptions of either non-ingressive or ingressive.

Although some instances of *gan* and *can* in ME are characterised by obscurity in meaning, as reported by the sources outlined in the previous paragraph, the opacity-like situation has been found to apply to some instances of *beginnen* in ME and Scots, too, even though the main function of this verb is to express ingressive aspect. As already mentioned in Section 5.2.2, some sources see this verb, especially when it

occurs with the plain infinitive, but also the *to*-infinitive, as occasionally interchangeable with *gan* and *can* in ME, thus implying either non-ingressive or ingressive meaning of *beginnen* (*MED* s.v. *biginnen* v. 6, Häusermann 1930: 21-22, Brinton 1981: 185-186, cf. Los 2000; 2005 on OE *beginnan*). But, only Visser (1969: 1580) points out that there are such cases of *beginnen* with infinitive where it may not be easy to precisely identify the meaning of this verb, “it is extremely difficult to make out whether or not, in a particular case, the verb *beginnen* has ceased to refer to the initial stage, the commencement, of the action”. Furthermore, Visser (1969) also provides an example of where *gan*, *did* and *began* appear to be similarly used in “successive versions” of *Cursor Mundi*, as in (41) below. This implies that the above verbs have the same or similar meaning and that they can be used in free variation, as suggested by the choice of forms in different MSS, but this does not necessarily suggest that the verb *beginnen* is subject to further grammaticalization, as changes in meaning can occur without grammaticalization taking place. However, although some instances of *beginnen* may be speculated to have a non-ingressive meaning, too, Visser (1969: 1580) concludes, “[t]he subjoined evidence is, admittedly, of a very tentative character, contextual clues for determining the ‘correct’ interpretation being either missing or inconclusive”. This emphasises the role of contextual information in an attempt to establish the meaning of this verb.

- (41) *and* *þus* *bi-gan* (*gan*, *began*, *dud* – different MSS)
 and *thus* *did/began*
 til *him* *to* *say*
 to *him* *to* *say*
 (*Cursor Mundi* 12296; Visser 1969: 1580)

‘and did/began to say to him in this manner’

An analysis of the findings of this study shows that in line with descriptions in the literature, not all meanings of *gan* and *can* in the NME and ESc data sets neatly situate themselves within the categories of either non-ingressive or ingressive. A considerable number of such instances can be found with *gan* in the NME data set, although some instances of *can* characterised by ambiguity, as far as their meaning is concerned, occur in the NME and ESc data sets, too. The main reason, it seems, for

this kind of categorisation is that some verbs found in infinitive with *gan* and *can* can be classed as either achievement or activity verbs, thus permitting either ingressive or non-ingressive interpretations of this verb and its variant. For example, *say* and *sai* used with *gan* and *can*, as in (42) below and (43) below, respectively, as well as *tell* found with *gan*, can be viewed as being with or without duration, referring to an act of uttering words or providing a description. Furthermore, the use of *ilkan* ‘each’ as the subject of *can*, as in (43) below, may facilitate an interpretation of the verb in infinitive as denoting a punctual event or situation, when the idea of saying is considered in isolation, or an iterative event or action, since more one than person is involved. Semantic opacity appears to characterise four instances of *beginnen* in the NME data set in Period II, too. For example, the verb *lose* (being prototypically punctual), as in (44) below and (45) below, may refer to a punctual event or situation that is non-repeatable, where the meaning of *beginnen* can be interpreted as non-ingressive. But, it may also denote the process of losing something that has been had and possessed (in both a metaphorical and non-metaphorical way), thus involving duration, in which case, the interpretation of *beginnen* would be ingressive. Furthermore, contextual information appears to support both interpretations here, too.

- (42) *Pan Peter þus to him gan say,*
 Then Peter thus to him did/began to say,
“Tel till vs, Lord, We þe pray
“Tell to us, Lord, We thee pray
What þir maters er to mene.”
 What these matters are to mean.”
 (1350-1420 *The Northern Homily Cycle* P II pp. 73-74; *HCET*)

‘Then Peter did/began to say to him in this manner, “Tell us, Lord, we beg you what matters/issues are to mean.”’

- (43) *And ilkan can til oþer sai,*
 And each did/began to to other say,
“Quat a sing es þis!” coth þai
 “What a sign is this!” said they
 (1350-1420 *Cursor Mundi* p. 1015; *HCET*)

‘And each did/began to say to the other, “What a sign is this!” they said’

(44) *and be þat at I sulde begyn to*
 and by that that I should do/begin to
lose þe vertue of my mekenes
 lose the virtue of my humility
 (1420-1500 *Alphabet of Tales* p. 509; *ICoMEP*)

‘and by that, that I would do/begin to lose the virtue of my humility’

(45) *Cesarius tellis how on a tyme ij men*
 Cesarius tells how on a time ij men
played at þe dyce, and when þe tane of
 played at the dice, and when the one of
þaim began to lose, he began to wax
 them did/began to lose, he began to grow
wrathe with þe toder & speke grete wurdis
 angry with the other & speak great words
 (1420-1500 *Alphabet of Tales* p. 83; *ICoMEP*)

‘Cesarius tells how at a time two men played the game of dice, and when one of them did/began to lose, he gradually began to grow angry with the other and spoke great words’

In addition to the problem of the verb in infinitive being interpreted as either an achievement or activity verb, verbs like *dyng*, as in (46) below, and *byte*, as in (47) below, both found with **gan** in the NME data set, pose a problem in relation to the classification of their internal temporal structure of the event or situation which these verbs denote. Both verbs can be described in terms of a single stage of their event or situation, without having clearly identifiable initial and final points. They can, therefore, be classified as semelfactive verbs. In the case of *dyng*, the actions of beating of Jesus will typically require a series of strikes, without the focus on either initial or final stages of this. However, it is also possible to view striking as complete in that it must have finished before Jesus was swung at the pillar, in which case the interpretation of **gan** would be non-ingressive. A similar interpretation can be offered for the verb *byte*. Eating an apple can be viewed as an action involving bites, which is different from that of eating, for example, a blueberry, which would not require a series of bites. However, biting an apple can be perceived here as complete, especially that in *The Holy Bible: King James Version* (2011: 2), the central aspect of this story illustrates the act of Adam and Eve consuming the fruit from the tree of knowledge of good and evil in the garden of Eden, rather than the action of eating it,

“[the woman] took of the fruit thereof, and did eat, and gave also unto her husband with her; and he did eat”.

- (46) *Ful fast þai gan hym dyng,*
 Full fast they did/began to him strike,
and at þe pyler swyng,
 and at the pillar swing,
& his fayre face defowlyng /
 & his fair face defouling /
with spittyng. /
 with spitting. /
 (1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 57;
ICoMEP)

‘They quickly did/began to strike him and swung him at the pillar, and defiled his beautiful face with spitting’

- (47) *Adam, thyselffe made al þis syte,*
 Adam, thyself made all this grief,
For to the tree þou wente full tyte
 For to the tree thou went full quickly
And boldelyon the frute gan byte
 And boldly on the fruit did/began to bite
My lord forbed
 My lord forbade
 (1420-1500 *The York Plays* p. 70; *HCET*)

‘Adam, you made all this grief yourself, because you very quickly went to the tree and boldly did/began to bite on the fruit that my lord forbade’

A number of verbs found with *gan* and *can* in the NME data set can in isolation be classed as activity verbs, i.e. verbs whose Aktionsart involves duration, i.e. a process that continues in time without an inherent goal (cf. Vendler 1967: 28) or stages (cf. Rothstein 2008: 18), or states. Yet, taking into account contextual information may result in a reinterpretation of these verbs and their subsequent reclassification to a different Aktionsart category altogether. This, in turn, may have an effect on the meaning of *gan* and *can* in the data sets under investigation. For example, the verb *striue* ‘strive’ in (48) below denotes an activity, in this case of quarrelling or wrangling with *þe fende* ‘the enemy’, with the likely interpretation of *gan* given as ingressive. However, the contextual information shows that this construction occurs

in conjunction with *gert*, which has the meaning of ‘to do, perform; to make’ (*OED* s.v. *gar* v. 1), and so this could imply a non-ingressive interpretation of *gan* in this example in that each of the events or situations presented here is complete. In the case of *wepe* ‘weep’ in (49) below, there might be a shift from activity to achievement, given the presence of the pronoun with the object NP *sekenes* ‘sickness’: thus, ‘she wept for sickness’ (activity) is different in meaning from ‘she wept for her sickness’ (achievement) in that the latter becomes more specific. On the other hand, the context suggests that an ingressive interpretation is also possible: the woman is aware of her sickness, and the Pharisee is not (*Scho knew hir sekenes, and he noght*); after the leech is brought between both of them (*ðe leche bitwene twa seke was broght*), the woman begins crying, admitting the fact that she is sick. Next is the verb *knele* in (50) below, which can refer to kneeling (activity), falling to the knees (accomplishment), or genuflecting (achievement), whereas with regards to *like* in (51) below, this verb can be seen as either a state verb, but also an achievement verb, resulting in different aspectual interpretations of *gan*, especially that the context does not provide additional information here.

- (48) *Also þis man þat þusgat died*
 As this man that thusgate died
By man-kind may be signifide
 By man-kind may be signified
First when þe fende gan with him striue
 First when the enemy did/began to with him strive
And gert him leue þe way of liue
 And made him live the way of life
 (1350-1420 *The Northern Homily Cycle* P III p. 137; *HCET*)

‘As the death of this man for mankind may be a first sign of when the enemy did/began to fight against him, and made him live the way of life’

- (49) *De woman to hirself toke kepe;*
 The woman to herself took keep;
For hir sekenes sare gan scho wepe
 For her sickness sore did/began to she weep
 (1350-1420 *The Northern Homily Cycle* P III p. 128; *HCET*)

‘The woman took notice; she did/began to weep greatly for her sickness’

- (50) *Pat it was iewus þan wist i wel,
That it was Jesus then knew I well,
And cuthli for him can i knele,
And certainly for him did/began to I kneel,
And i said, ‘ai laurd! blisced þou be
And I said, ‘Aye Lord! Blessed you be
(1350-1420 *Cursor Mundi* p. 1014; *HCET*)*

‘I knew well that it was Jesus, and certainly I did/began to genuflect/kneel before him, saying, ‘Yes, Lord! Blessed be you’

- (51) *Sen þat þe lest [{blis}] was slike,
Seeing that the active bliss was such,
þat þe monk so wele gan like
That the monk so well did/began to like
þat of thre hundreth zeres and ma
that of three hundred years and more
(1350-1420 *The Northern Homily Cycle* P II p. 206; *HCET*)*

‘Seeing that the active bliss was such, which the monk so well did/began to like, that of three hundred years and more’

In summary, known to occur in the process of grammaticalization, opacity refers to a situation where two interpretations are possible. It can also be seen in relation to *gan* and *can*, as well as some isolated instances of *beginnen*, in the NME and ESc data sets, resulting in more than one interpretation of these verbs in some contexts.

5.4.3 Persistence

Apart from semantic bleaching and opacity, a number of instances of *gan* in the NME data set are characterised by persistence. Their occurrence, though, is also another indication that the verb *gan* grammaticalizes as an auxiliary in this data set. Hopper (1991: 22) argues about persistence, “[w]hen a form undergoes grammaticalization from a lexical to a grammatical function, so long as it is grammatically viable some traces of its original meanings tend to adhere to it, and details of its lexical history may be reflected in constraints on its grammatical distribution” (cf. Hopper & Traugott [1993] 2003: 124). He also notes about persistence that “[the relationship between the meaning and function of a grammatical construction and its etymological source] is often completely opaque by the stage of morphologization, but during intermediate stages it may be expected that

a form will be polysemous, and that one or more of its meanings will reflect a dominant earlier meaning” (Hopper 1991: 28). This should be understood that although the semantically bleached meaning of *gan* and *can* is not a continuation of the original lexical meaning of OE *onginnan* and OE and eME *aginnan*, it is still possible to observe polysemy, namely the presence of instances of *gan* with the ingressive meaning in the NME data set. The occurrence of such instances, however, would also be indicative of the degree of grammaticalization of an item.

As is the case with other instances of *gan* and *can* in the NME and ESc data sets, the ingressive meaning of *gan* is the result of the Aktionsart of the verb in infinitive, although the role of contextual information plays a crucial role in aspectual interpretations here, too. For example, *glyde*, as in (52) below, could be classified as belonging to Vendler’s (1957; 1967) category of activity verbs and so, the action described here refers to the movement of the streams of Jesus’ blood, with the AdvP *downe* ‘down’ indicating the direction or path, this, therefore, being a dynamic activity. In this case, *gan* would be seen as used in an ingressive sense, especially that the action of Jesus’ blood flowing is compared to the streams of a river, which in itself denotes a continuous activity. Another instance of *gan* characterised by persistence involves the activity verb *playe*, as in (53) below, where the focus is on the male infant sucking the lioness’ milk and, subsequently, beginning to play with her, both events presented here in succession. More specifically, after the feeding of the child is finished, presented here as an activity, rather than an event, the infant begins playing with the lioness, presented to the reader as an incomplete activity. During the activity of playing, which now becomes a background activity, the male infant is seen as going back to sucking the lioness’ breast, which can still be perceived as an on-going activity of playing, but an interrupted one. The AdvP *Umwhile* ‘Sometimes’ signals here the cause of this interruption of the event or action, implying that the event or action of playing is not finished yet, whereas the reference to kissing and embracing constitutes a separate event or situation here altogether.

(52) *In five steads of his flesch þe blode gan*
 In five steads of his flesh the blood began

downe glyde, als stremes of þe strande
 down glide, as streams of the river
 (1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 57;
ICoMEP)

‘In five parts of his body, the blood began to glide down, like streams of the river’

(53) *The lyones fonde thay in hir dene,*
 The lioness found they in her vale,
A knave childe laye sowkand hir then
 A male infant lay sucking her then
And gan with the lyones to playe.
 And began with the lioness to play.
Umwhile the childe sowkede hir pappe,
 Sometimes the child sucked her breast,
Umwhile gan thay kysse and clappe;
 Afterwards did they kiss and embrace;
 (1420-1500 *Octavian* l. 442-446; *TEAMS*)

‘They [the men] found the lioness in her vale, a male infant lay sucking her then and began to play with her. Sometimes the child sucked her breast, afterwards, they kissed and embraced;’

In sum, although the majority of instances of *gan* and *can* are non-ingressive in the NME and ESc data sets, there is evidence to show that traces of the original ingressive meaning can still be found amongst instances of *gan* in the NME data set.

5.5 Summary

This chapter started off with an overview of different types of meaning found with *gan* and *can* in ME and Scots. This verb and its variant are mainly used with non-ingressive meaning, but a number of sources also show that *gan* and *can* can be used with ingressive sense, in addition to instances of this verb and its variant being difficult to classify. This overview also showed that the verb *beginnen* is used with ingressive meaning in ME and Scots, although the *MED* (s.v. *biginnen* v. 6) reports instances of this verb in a non-ingressive sense, too, in addition to Los (2000; 2005) arguing that OE *beginnan* with the plain infinitive shows bleached uses, too, in Ælfric. My findings demonstrated that the meaning of *gan* could be categorised as non-ingressive, ingressive and ambiguous in the NME data set, whereas that of *can* could be classed as non-ingressive and ambiguous only in the NME and ESc data

sets. The verb *beginnen*, on the other hand, expresses ingressiveness in the NME and ESc data sets (this also applies to the variant form *begouth* in the ESc data set), although a small number of instances of this verb could be interpreted as ambiguous in the NME data set. I then presented a literature review on how other sources have determined the meaning of *gan* and *can* in ME and Scots. The analysis of instances of *gan* and *can*, as well as *beginnen* in the NME and ESc data sets demonstrated that the Aktionsart of the verb in infinitive, adverbials of duration or iteration, but also contextual information can have an effect on the meaning of these verbs. It showed that instances of this verb and its variant are semantically bleached these data sets, but where two interpretations are possible, this is mainly due to opacity. Only a small number of instances of *gan* are considered to show persistence in the NME data set. Overall, therefore, the analysis of the findings of this chapter strengthens the claim that *gan* and *can* are in the process of grammaticalization in the NME and ESc data sets to gradually become associated with non-ingressive meaning, focussing on the completion of an event or situation denoted by the verb in infinitive. The verb *beginnen*, on the other hand, shows very little change in terms of its meaning, being associated with expression of ingressiveness in the NME and ESc data sets. However, in line with the path of grammaticalization, this meaning change of *gan* and *can* follows changes in context, as discussed in the following chapter.

Chapter Six – Changes in Context

6.1 Introduction

So far, I have investigated the auxiliations of *gan* and *can* in the NME and ESc data sets from the perspective of change that this verb and its variant undergo from lexical to grammatical. In this chapter, I consider the grammaticalization of *gan* and *can* in the NME and ESc data sets from the pragmatic point of view. Since instances of *gan* and *can* predominantly occur in verse, a metrical function has been proposed for this verb and its variant: to provide an extra syllable in a line of verse (Taylor 1917: 574, Beschorner 1920: 15-19, Funke 1922: 22, Koziol 1932: 132-133, Kerkhof 1966: 154, Smyser 1967: 68-69, Visser 1969: 1572, Terasawa 1974: 102, Tajima 1975: 434, 437-438, Mustanoja 1983, Fischer 1992: 266, Richardson 1991a: 80-81, Ogura 1998: 300, Sims 2008: 140, Fulk 2012: 106-107) and/or to enable the verb in infinitive to provide rhyme (Koziol 1932: 132-133, Kerkhof 1966: 154, 158, Smyser 1967: 74, Terasawa 1974: 102, Tajima 1975: 434, 437-438, Brinton 1981: 164). Other proposals on a role of *gan* and *can* in verse include: an intensive-descriptive or stylistic function (Wuth 1915, Funke 1922: 8, Häusermann 1930: 18-22, Homann 1954: 389-398, Mustanoja 1960: 611-614, Kerkhof 1966: 154, 156, Terasawa 1974: 101-104), although now shown to be subjective and difficult to prove, and a textual one (Brinton 1988; 1996), especially that *gan* has a “skewed” distribution in ME texts and it is associated with “salient and important turns in the course of events”, “occur[ing] in multiples, clustering around significant events” (Brinton 1996: 77). Therefore, while the metrical function can explain the invariant aspects of *gan* and *can* in ME and Scots verse, the stylistic and textual ones can account for the variable aspects of this verb and its variant, in line with principles of semantic and pragmatic change in the process of grammaticalization (cf. Brinton 1996: 82).

In Section 6.2, I provide an overview of the context associated with the occurrence of *gan* and *can* in ME and Scots verse and discussed in the literature on the subject. In this section, I also present my findings, namely the types of texts and linguistic environment in which this verb and its variant occur (the information presented there includes findings on meter and on rhyme patterns of the verse texts from the NME

and ESc data sets in which *gan* and *can* feature). In Section 6.3, I review proposals on various functions of *gan* and *can* that have been proposed in the literature: metrical and rhyming, intensive-descriptive and textual. In this section, I also provide an analysis of my findings, where I show that *gan* and *can* are an integral part of verse in the NME and ESc data sets – exclusively the octosyllabic line in Period I, but also other types of line in Period II – with this verb and its variant also making the verb in infinitive available to provide rhyme. In addition, I argue that *gan* and *can* have a textual function, in which they isolate events or situations and changes in behaviour of some characters from the rest of the narrative, with an interpersonal function focussing on the importance of such events or situations. Since the verb *beginnen* is almost exclusively associated with expression of ingressive aspect (Chapter Five), and the findings of this chapter do not show anything unusual in terms of this verb's occurrence in the textual material under investigation (linguistic context, metre and rhyme), I mainly concentrate here on *gan* and *can*, with references made to *beginnen* being for comparison only.

6.2 Context of *Gan* and *Can* in Middle English and Scots

In the following sections, I investigate the context (i.e. text type and its associated metre and rhyme, and linguistic environment) in which the verb *gan* and its variant *can* are described to occur in ME and Scots. I also show that there is a relationship between *gan* and *can* in the NME and ESc data sets and the type of context in which this verb and its variant occur: a) *gan* and *can* in these data sets are restricted to verse, which is mainly classed as octosyllabic in Period I, with this verb and its variant being found in other types of verse, too, but in the NME data set and in Period II only; b) the occurrence of *gan* and *can* appears to be restricted to foregrounded clauses, which are declarative (cf. Terasawa 1974: 102, Brinton 1996: 76); and c) the verb in infinitive occurring with *gan* and *can* is predominantly or exclusively found in rhyme position in the NME and ESc data sets. Let us firstly turn to the type of text and linguistic environment in which the verb *gan* and its variant *can* are found in ME and Scots, including the NME and ESc data sets.

6.2.1 Text and Linguistic Environment

Both *gan* and *can* occur in verse in ME and Scots (Taylor 1917: 574, 590, Beschorner 1920: 15-16, 17-19, Funke 1922: 22, Koziol 1932: 132-133, Smyser 1967: 68-69, 74, Tajima 1975, Brinton 1981: 164; 1983: 236; 1996: 69, Fischer 1992: 266, Ogura 1998: 300, Sims 2008: 140, cf. Mossé 1952: 103, Mustanoja 1960: 610-615). In this type of writing, as opposed to in prose, the author imposes a greater degree of structure. Although most prose can be classed as narrative, the general principles of narrative usually apply to prose, with very little narrative verse. However, the types of texts in which *gan* and *can* are most frequently found are, in fact, narrative verse texts or “rhymed romances” (Brinton 1996: 69). In relation to ME, the following verse texts amongst others can, therefore, be found in the literature as containing instances of this verb and its variant: *Ormulum* (Sims 2008: 141-144), *Sir Gawain and the Green Knight*, *Pearl*, *Purity and Patience* (Koziol 1932: 132-133, Homann 1954: 394, Tajima 1975: 429, 432), Chaucer’s works, although to a varying degree (Funke 1922: 26, Smyser 1967: 68-69, Homann 1954: 394, Terasawa 1974: 100, Brinton 1983: 236; 1996: 75-78, Richardson 1991a: 165) and rhymed romances, such as *Amis and Amiloun*, *Athleston*, *Lazamon’s Brut*, *Carle of Carelyle*, *Confessio Amantis*, *Degare*, *Degrevant*, *Sir Eglamour of Artois*, *Emare*, *Floris and Blancheflour*, *Gamelyn*, *Gawen and Dame Ragnell*, *Havelok the Dane*, *King Horn*, *Lai le Freine*, *Launfal*, *Le Morte Arthur*, *Libeaus Desconus*, *Low Degree*, *Sir Orfeo*, *Richard Coer de Lyon*, *Tristrem*, and *Ywain and Gawain* (Terasawa 1974: 100).

The list of texts with instances of *gan* and *can* also includes a number of NME and ESc works. In relation to northern texts, Ogura (1997: 408), for example, shows that this verb and its variant occur in *Cursor Mundi*, while the *MED* (s.vv. *can* v.; *ginnen* v. 3b) and the *OED* (s.vv. *can* v.² I 1a; *gin* v.¹ 1) list instances of this verb or its variant from *The Prick of Conscience*, *Sir Isumbras* and *The York Plays*. Terasawa (1974: 94-96) provides frequencies of occurrence of the “*gan* idiom [*gan* and *can*]” in northern texts like *Awntyrs*, *Eglamour*, *Tottenham* and *Tristrem*, but his classification is based on the “Dialect of Original Composition”, rather than the dialect of an MS, which may not display an accurate picture of the distribution of the

“*gan* idiom” in northern texts. In relation to ESc works, Macafee (1992/93: §8.13) argues, “*gan* [and *can*, as well as *couth*] are ... used in more native styles”. Such styles include Lydgate’s *The Troy Book*⁹⁹ and *King Hart* (DSL s.v. *gan* p.t.), *Golagrus* and Barbour’s *The Bruce* (Terasawa 1974: 96). The variant ***can***, on the other hand, occurs in *The Scottish Legendary*, Wynthoun’s *The Original Chronicle of Scotland*, Blind Harry’s *Wallace*, Holland’s *The Book of the Howlat* and Henryson’s *The Poems of Robert Henryson* (DSL s.v. *can* v. 2, Taylor 1917: 574, 579), while ***couth*** in *The Scottish Legendary*, Blind Harry’s *Wallace*, Holland’s *The Book of the Howlat*, Henryson’s *The Poems of Robert Henryson* and Rolland’s *The Book of Seven Sages* (DSL s.v. *can* v. 2). However, between different MSS containing Barbour’s *The Bruce*, the verb ***gan*** solely occurs in one, whereas the variant ***can*** in the other (DSL s.vv. *can* v. 2; *gan* p.t.).

Although ***gan*** and ***can*** are a feature of verse, some sources report finding occasional instances of ***gan*** in prose, too. For example, Brinton (1981: 164) claims, “in Chaucer’s prose, gan appears only three times in the ‘Tale of Melibeus’, always with inchoative meaning, and not at all in his other prose” (cf. Smyser 1967: 69, Brinton 1988: 120; 1996: 69). However, the use of ***gan*** is not restricted to Chaucer. Brinton (1996: 69), for example, maintains that in the *HCET*, “only 8% [of ***gan***] occur in prose, and there is only one example [of this verb] in an alliterative poem[, too]”. Nevertheless, apart from providing such statements, sources like Brinton (1981; 1996) offer very little information about the distribution of ***gan*** in this mode of writing. In particular, they provide no detail about the context in which such instances of ***gan*** are found in prose works, even though such detail would be crucial to our understanding of a function or functions of ***gan***, and ***can***, in ME and Scots. Let us now turn to the distribution of ***gan*** and ***can***, as well as ***beginnen*** in terms of literary form in the NME and ESc data sets. My findings show that ***gan*** and ***can*** occur in verse in the NME and ESc data sets, as outlined in Table 50 and Table 51,

⁹⁹ Lydgate’s *The Troy Book* contains Scottish fragments (the DSL’s edition of *The Troy Book* is from Horstmann (1881-1882) and listed as authored by John Barbour). On the basis that this work contains the lines, “Her endis the monk and begynnys barbour” and “Her endis barbour and begynnys the monk”, Horstmann (1881-1882) concluded that John Barbour was the author of *The Troy Book* fragments, i.e. “sections of a translation of Guido delle Colonne’s Latin history of the taking of Troy (Duncan 2004). However, this authorship has now been shown to be questionable (cf. von Contzen 2016: Appendix: The Scottish Legendary: authorship, dialect, and arrangement, Duncan 2004).

respectively, with two instances of **gan** recorded in a text classed as prose in the NME data set, namely Rolle's *Richard Rolle of Hampole, Yorkshire Writers* (the distribution of the verb **beginnen** in different types of text in the NME and ESc data sets is provided for comparison in Table 52 and Table 53, respectively). This, therefore, matches descriptions in the literature on the distribution of this verb and its variant in ME and Scots, although I show further below that the two instances of **gan** occurring in what is classed by the authors of the *ICoMEP* as a prose text, in fact, occur in a passage that has verse characteristics.

Table 50

Distribution of gan and can with Infinitive in Types of Text in the NME Data Set in Period I and Period II

	gan						can		
	Verse			Prose			Verse		
	<i>N</i>	<i>R_I</i>	%	<i>N</i>	<i>R_I</i>	%	<i>N</i>	<i>R_I</i>	%
Period I	15	14	88	2	2	12	8	7	100
Period II	75	26	100	-	-	-	12	4	100

Notes: *N* – Observed absolute frequency
R_I – Observed relative frequency per period per 100,000 words
 % – Per-cent of **gan** and **can** per type of text
 NB: Observed relative frequency is rounded to the nearest integer

Table 51

Distribution of can with Infinitive in Types of Text in the ESc Data Set in Period I

	can		
	Verse		
	<i>N</i>	<i>R_I</i>	%
Period I	8	24	100

Notes: *N* – Observed absolute frequency
R_I – Observed relative frequency per period per 100,000 words

% – Per-cent of **can** per type of text
 NB: Observed relative frequency is rounded to the nearest integer

Table 52

*Distribution of **beginnen** as an Independent Verb and with Infinitive in Types of Text in the NME Data Set in Period I and Period II*

	<i>beginnen</i>											
	Other						Inf					
	Verse			Prose			Verse			Prose		
	<i>N</i>	<i>R₁</i>	%	<i>N</i>	<i>R₁</i>	%	<i>N</i>	<i>R₁</i>	%	<i>N</i>	<i>R₁</i>	%
Period I	5	5	21	19	17	79	12	11	44	15	14	56
Period II	9	3	24	29	10	76	16	6	12	119	42	88

Notes: *N* – Observed absolute frequency
R₁ – Observed relative frequency per period per 100,000 words
 % – Per-cent of **beginnen** per type of text
 NB: Observed relative frequency is rounded to the nearest integer

Table 53

*Distribution of **beginnen** as an Independent Verb and with Infinitive in Types of Text in the ESc Data Set in Period I and Period II*

	<i>beginnen</i>											
	Other						Inf					
	Verse			Prose			Verse			Prose		
	<i>N</i>	<i>R₁</i>	%	<i>N</i>	<i>R₁</i>	%	<i>N</i>	<i>R₁</i>	%	<i>N</i>	<i>R₁</i>	%
Period I	-	-	-	7	21	100	1	3	100	-	-	-
Period II	-	-	-	55	12	100	-	-	-	11	2	100

Notes: *N* – Observed absolute frequency
R₁ – Observed relative frequency per period per 100,000 words
 % – Per-cent of **beginnen** per type of text
 NB: Observed relative frequency is rounded to the nearest integer

Although collectively categorised as verse, such texts in the NME and ESc data sets differ in terms of genre labels. I now turn to then categories of verse texts in these data sets where this verb and its variant are found. As explained in Section 2.2.3.3, in order to remain consistent and to avoid misclassification, overlapping and bias, I maintain genre labels used by the compilers of the corpora used in this study. The verb *gan* and its variant *can* are, therefore, found in verse texts or their extracts classified as: ‘romance’ (*Sir Eglamour of Artois* and *Octavian*), as in (1) below, ‘homily’ (*The Northern Homily Cycle*), as in (2) below, ‘historical account’ (*Cursor Mundi*), as in (3) below, ‘legend’ (*The Life of St. Julian Hospitaller* and *St. Andrew and the Three Questions*), as in (4) below and (5) below, ‘religious treatise’ (*The Pricke of Conscience*), as in (6) below, ‘drama mystery’ (*The York Plays*), as in (7) below. When considering all of these categories of text together, in which *gan* and *can* occur, these texts could also perhaps be described as having a religious, didactic and historical function, falling into the type of ‘serious literature’. However, literary works on medicine, science and philosophy could also be fall into the category of didactic. For example, *The Liber de Diversis Medicinis*, classified as ‘handbook medicine’ by the compilers of the *HCET*, as part of the NME data set, also has a didactic function – it contains a collection of medical advice and recipes. Nevertheless, what generally distinguishes the above genre types from that of *The Liber de Diversis Medicinis* is that they contain a narrative (cf. Brinton 1996: 69), even though they would still be seen as works of instruction, at the same time raising a question about the classification of literary works.

- (1) *Sche ys so bryght of ble!”*
 She is so fair of face!”
 The knyght gan the Erle pray,
 The knight did/began to the Earl beg,
 “*Lord, I have served the many a day*
 “*Lord, I have served thee many a day*
 Wouche ye her ave on me.”
 Vouch ye her always on me.”
 (1420-1500 *Sir Eglamour of Artois* l. 208-213; *TEAMS*)

‘She is so beautiful!’ The knight did/began to beg the earl, “Lord, I have served many a day, betroth her to me.”

- (2) *Pe woman þan hir dedes gan rew*
 The woman then her deeds did repent
And þus scho said vnto Jesu,
 And thus she said unto Jesus,
 (1350-1420 *The Northern Homily Cycle* P II p. 82; *HCET*)

‘The woman then repented her sins and said to Jesus in this way,’

- (3) *Pe slogth he fand þat him gan wiss*
 The path he found that him did mark
Tilward þe gate of paradis;
 Toward the gate of paradise;
 (1350-1420 *Cursor Mundi* p. 82; *HCET*)

‘He found the path that marked the way for him towards the gate of paradise;’

- (4) *A new dysese thane can he tak,*
 A new hardship then did he suffer,
Seand his wyf sic sorou mak;
 Seeing his wife such sorrow make;
 (1375-1420 *The Life of St. Julian Hospitaller* l. 386-390; *TEAMS*)

‘He then suffered new hardship, seeing that his wife brought so much sorrow;’

- (5) *Sad to hyme agane, “Thu ga*
 Said to him again, “Thou go
Til hyre that cane this demand ma
 To hir that did this demand make
And spere at hyre grathly.
 And ask at her promptly.
 (1375-1420 *St. Andrew and the Three Questions* l. 1084-1087;
TEAMS)

‘[the porter] said to him again, “Go to her that made this demand and ask her immediately.”’

- (6) *And say “weles us þat here er wonnand”,*
 And say “choose us that here are living”,
And thank God omang þat þam gan wysse
 And thank God together that them did point
Til mekenes, þat þam led til þat blysse
 To meekness, that them led to that bliss
 (1350-1420 *The Pricke of Conscience* p. 250; *HCET*)

‘and said, “choose us that are living here”, and together thanked God, who pointed them to meekness, and led them to that bliss’

(7) Adam, thyselffe made al þis syte,
Adam, thyself made all this grief,
For to the tree þou wente full tyte
For to the tree thou went full quickly
And boldelyon the frute gan byte
And boldly on the fruit did/began to bite
My lord forbed
My lord forbade
(1420-1500 *The York Plays* p. 70; *HCET*)

‘Adam, you made all this grief yourself, because you very quickly went to the tree and boldly did/began to bite on the fruit that my lord forbade’

As already alluded to earlier, the prose work by Rolle titled *Richard Rolle of Hampole, Yorkshire Writers* in the *ICoMEP*, which is part of the NME data set used in this study, contains two instances of **gan**. This text falls into the category of ‘religious treatises’, which can be associated with religious and didactic literature, too. However, the context where these instances occur is different from that of prose where the verb **beginnen** is, for example, found. They occur in *Meditatio de passione Cristi* found in the above prose sample, with this extract displaying some verse-like characteristics, such as metre and rhyme, even though it is written as prose in the MS containing this text (these verse-like features of this part of *Richard Rolle of Hampole, Yorkshire Writers* are discussed in depth in Sections 6.2.2 and 6.2.3). Furthermore, this extract is a narrative depicting the Passion of Christ. The use of such a passage with verse-like characteristics in a prose work could perhaps be explained by the fact that insets like this one, even though they are not clearly marked off the rest of the text, were meant to extract specifically memorable passages from longer works (Whitehead 2005: 111) to create a graphic image in the listener’s or reader’s mind, in this case, about Christ’s Passion. Finally, the occurrence of such mixed forms, i.e. verse vs. prose, would typically depend on text type, genre, and a degree of formality, with Hiatt (2007: 291) arguing that “[m]ixing genres ... may turn out to be the fundamental trait of Middle English literature”.

Given the fact that *Meditatio de passione Cristi* is characterised by verse-like properties, I have made an attempt to re-arrange its layout to identify the position of

the verb in infinitive found with *gan*, given that one of the functions quoted in the literature of this verb is to enable the verb in infinitive to provide rhyme in a line of verse (Koziol 1932: 132-133, Kerkhof 1966: 154, 158, Smyser 1967: 74, Terasawa 1974: 102, Tajima 1975: 434, 437-438, Brinton 1981: 164) (cf. Appendix Two and an edition of Cambridge University MS DD.v.64 Part 3 by Horstman 1895 and the way *Meditatio de passione Cristi* is presented there¹⁰⁰) (metre, and internal and end rhyme of this extract are discussed in Sections 6.2.2 and 6.2.3, respectively). To do this, I used other similar verse extracts presented as prose at the end of the above MS, with the exception that they have distinct markings to denote couplets, tercets and quatrains, with stanzas containing the other instances of *gan* organised into a quatrain, although there is a possibility that these are added at a later date (cf. Appendix Three).¹⁰¹ However, regardless of whether the authorial amendments seen there are contemporary to the date of the MS or not, dividing *Meditatio de passione Cristi* into stanzas proved difficult. As a result, I have divided this piece into two clearly distinct parts, although each still subdivided into couplets, tercets and quatrains, probably thematically corresponding to different parts of the description of the Passion of Christ. These two parts are characteristic of a prayer, this being an example of genre mixing in medieval texts, too: a) the narrative or descriptive one, which contains the two instances of *gan* with infinitive, as in (8) below, and a supplication, as in (9) below. Overall, however, the occurrence of *gan* in this this passage shows that in the NME and ESc data sets, both *gan* and *can* are associated with verse only, even though textual evidence may be categorised otherwise.

- (8) // My keyng þat water grette, and blode swette;
 sythen ful sare bette, so þat hys blode hym wette,
 when þair scowrges mette. /

¹⁰⁰ Horstman (1895: 57, 60) provides two northern versions of the MS containing *Richard Rolle of Hampole, Yorkshire Writers*: Cambridge University MS DD.v.64 and MS Rawlinson A 389. With regards to the use of *gan* in *Meditatio de passione Cristi*, the former MS has <gan dyng> and <gan glyde>, whereas the latter <gun dyng> and <gan glyde>. The tract *Pe forme of liuyng*, which contains the above extract, is written “in a sort of rhythmical, cadenced prose, so that frequent alliterative verses can easily be made out; **it would, however, be fruitless to attempt transcribing it into verse**” (Horstman 1895: 3; emphasis mine). Although Horstman (1895) seems to refer here to the whole tract, rather than *Meditatio de passione Cristi* only, he suggests that the above extract has verse-like characteristics, too.

¹⁰¹ Incidentally, the poems at the end of Cambridge University MS DD.v.64 contain *gan* with *dyng* and *glyde*, too. However, these are excluded from this study, as they are not part of the *ICoMEP* corpus, constituting different instances from those found in *Meditatio de passione Cristi*.

Ful fast þai gan hym dyng, and at þe pyler swyng,
& his fayre face defowlyng |r[!] with spittyng. /
þe thorne crownes þe keyng, ful sare es þat prickyng.
Alas my ioy and my swetyng es demed to hyng! /

Nayled was his handes, nayled was hys fete,
& thyrlid was hys syde so semely & so swete. /

Naked es his whit breste, & rede es his bloody syde;
wan was his fayre hew, his wowndes depe & wyde.
In fyue stedes of his flesch þe blode gan downe glyde,
als stremes of þe strande, hys pyne es noght to hyde. /

Dis to see es grete pyte, how he es demed to þe dede,
and nayled on þe rode-tre, þe bryght aungels brede. /

Dryuen he was to dole þat es owre gastly gude,
and als_so in þe blys of heuen es al þe aungels fude.
A wonder it es to se, wha sa vnderstude,
how god of mageste was dyand on þe rude. /

Bot suth þan es it sayde þat lufe ledes þe ryng;
þat hym sa law hase layde, bot lufe it was na thyng. /

(9) Ihesu, receyue my hert, & to þi lufe me bryng:
al my desyre þou ert, bot I couete þi comyng. /

Pow make me clene of synne, & lat vs neuer twyn;
kyndel me fire with-in, þat I þi lufe may wyn,
and se þi face Ihesu in ioy þat neuer sal blyn. /

Ihesu, my saule þou mend, þi lufe in to me send,
þat I may with þe lend, in ioy with-owten end. /

In lufe þow wownde my thought, and lyft my hert to þe:
my sawle þou dere hase boght, þi lufer make it to be.
þe I couete, þis worlde noght, & for it I fle;
þou ert þat I haue soght: þi face when may I see? /

Pow make my sawle clere, for lufe chawnges my chere:
how lang sal I be here?[when mai I negh þe nere, þi melody to here,]/

Oft to here sang,
þat es lastand so lang?

Pou be my lufyng,
þat I [þi] lufe may syng.

Restrictions on the occurrence of *gan* and *can* do not appear to be limited to verse only, but also the context of language. For example, the distribution of this verb and its variant is shown to be in the following linguistic environments in the NME and ESc data sets: a) declarative clauses, not in interrogative, imperative or negative ones, which is also in line with the findings by Terasawa (1974: 102) and Brinton (1996: 76), although in relation to other texts; b) descriptions of the narrative character (cf. Brinton 1996: 76), with only two instances of *can* and none of *gan* occurring in direct speech, as in (10) below; c) clusters (cf. Brinton 1996: 77), especially in longer *Octavian*, as in (11) below, although this is not always the case in this work and, especially, in others; in fact, only three stanzas in *Octavian* from the *TEAMS* have either three or four instances of *gan* grouped together. However, the remaining 35 stanzas containing this verb consisting of either one or two instances of it each, although mostly only one, which is also what characterises the distribution of *gan* and *can* in the remaining verse texts in the NME and ESc data sets; and d) main participants (cf. Brinton 1996: 76), for example, in the case of *Octavian*, this is *Clement*, but also *The Sowdane* ‘The Sultan’, as in (11) below, in addition to the Empress, one of her children that is captured by the lioness, and the lioness herself, all participating in the story (the narrative). Overall, therefore, the linguistic environment in which *gan* and *can* are recorded in the NME and ESc data sets is somewhat restricted, especially when comparing the context of this verb and its variant to that of *beginnen*. For example, the verb *beginnen* can be found in all types of clauses identified above, narrative descriptions, including direct speech, although to a lesser extent, and with a wider range of subjects, including the weather *it*, as in (12) below, and the existential *there*, as in (13) below.

- (10) *Sad to hyme agane, “Thu ga*
Said to him again, “Thou go
Til hyre that cane this demand ma
To hir that did this demand make
And spere at hyre grathly.
And ask at her promptly.

(1375-1420 *St. Andrew and the Three Questions* l. 1084-1087;

TEAMS)

‘[the porter] said to him again, “Go to her that made this demand and ask her immediately.”’

- (11) *One the morne when the day was lyghte,*
On the morning when the day was light,
Clement gan hymselven dyghte
Clement did himself prepare
Lyke an unfrely fere
Like an repulsive companion
And went into the heythen oste
And went into the heathen host
Thore the presse was althermoste,
Where the crowd was uttermost,
A Sarazene als he were.
A Saracen also he were.
And to the paveleone he gan wynn
And to the tent he did go
There the Sowdan hymselfe was in.
Where the Sultan himself was in.
Full brymly he gan bere
Full boldly he did shout
And askede tham sum of thaire mete.
And asked them some of their food.
Full wele he couthe thaire speche speke;
Full well he knew their spech speak;
The Sowdane hymselfe gan here.
The Sultan himself did hear.
(1420-1500 *Octavian* l. 1555-1566; TEAMS)

‘On the morning, when the day was light, Clement prepared himself like an ugly companion and went into the heathen host where the crowd was greatest, he was also a Saracen. And he went to the tent in which the Sultan was. Boldly he then shouted and asked them for some of their food. He knew well the language that they spoke, as the Sultan heard himself.’

- (12) *It shall begyn full sone to rayn vncessantle’,*
It shall begin full soon to rain incessantly’,
After dayes seuen be done, and induyr dayes forty,
After days seven be done, and anotherdays forty,
(1420-1500 *The Wakefield Pageants in the Towneley Cycle* pp. 17-18; HCET)

‘It will begin to rain incessantly’, after seven days have passed, and another forty days,’

(13) *And with-in a few yeris after he was
 And within a few years after he was
 made monke, þer began to brede a sur
 made monk, there began to spread a sore
 like a sylk threde in his legg,
 like a silk thread in his leg,
 (1420-1500 *Alphabet of Tales* p. 332; *ICoMEP*)*

‘And within a few years after he was made a monk, an ulcer began to spread on his leg like a silk thread,’

The type of form exploited for narrative verse in ME and Scots was usually octosyllabic. Let us now turn to a description of metre of the contexts featuring *gan* and *can* in the NME and ESc data sets.

6.2.2 Metre and Rhythm

One of the functions of *gan* and *can* described in the literature is that this verb and its variant are a metrical device to provide an extra syllable to fill a line of verse (the function of this verb and its variant to enable the verb in infinitive to provide rhyme in a line of verse is discussed below) (Taylor 1917: 574, Beschorner 1920: 15-19, Funke 1922: 22, Koziol 1932: 132-133, Kerkhof 1966: 154, 158, Smyser 1967: 68-69, Visser 1969: 1572, Terasawa 1974: 102, Tajima 1975: 434, 437-438, Mustanoja 1983, Fischer 1992: 266, Richardson 1991a: 80-81, Ogura 1998: 300, Sims 2008: 140, Fulk 2012: 106-107, cf. Brinton 1996). Verse is concerned with the use of sound effects, namely metre and rhythm, and sound-patterning, in addition to lexical and syntactic effects (J. Smith 2012: 55). However, although it has been widely accepted that versification is interplay of both metre and rhythm, leading some to even treat both terms as synonymous and interchangeable (Strachan & Terry [2000] 2011: 72), it is important to make a terminological distinction here. Attridge (1995: 3, 7), for example, helpfully defines metre as “an organising principle which turns the general tendency toward regularity in rhythm into a strictly-patterned regularity, that can be counted and named”, whereas rhythm as “a patterning of energy simultaneously produced and perceived; a series of alternations of build-up and release, movement and counter-movement, tending toward regularity but complicated by constant variations and local inflections” (cf. also Bath & Furniss

[1996] 2007: Rhythm and Metre, Hobsbaum 1996: 1-9, Harding [1976] 2010: 17-30, Spinks 2010: 47, Strachan & Terry [2000] 2011: 72, Leech [1969] 2013: Rhythm and Metre and Mayor [1901] 2014: 1-11 amongst others). Therefore, the term ‘metre’ is usually confined to the study of poetry, whereas the term ‘rhythm’ has some significance outwith the formal constraints of verse.

Although metre and rhythm are two different, but related terms, various types of versification rely on both the ‘natural’ origins of rhythm and the ‘cultural’ tradition of metrical form (Attridge 2013: 124). For example, metres in Germanic languages were accentual. This means the presence of four strong stressed syllables per line, but a lack of metrical feet and a number of unstressed syllables. Under the influence of French, however, English-speaking poets in ME have used accentual-syllabic metre, offering a compromise between stress and syllable counting. Therefore, the accentual-syllabic tradition of versification at the time combined two principles: a) the strict syllable counting of the Old French (OF) and Anglo-Norman *vers octosyllable*; and b) stress alternations characterising English speech (Minkova 2007: 186). However, since the accentual-syllabic metre combined two different ways of versification, diversity characterised this new way of versifying in ME. Attridge (2013: 124, emphasis in the original), for example, argues, “[b]y controlling both the number of syllables *and* the number of stresses in a metrical unit, it brings the two rhythmic resources into harmony; but it’s an unstable accord, and the strength and the variety of accentual-syllabic metre spring from its exploitation of that instability.” The emergence of the accentual-syllabic versification in ME was, therefore, about innovation.

Following the abandonment of the structural aspects of OE alliterative versification, a vast amount of 14th century literature was reinvented, but also composed, in the tradition of alliterative verse, resulting in works like *Sir Gawain and the Green Knight*, *Piers Plowman* and *The Alliterative Morte Arthure*. However, at the same time, new forms of versification, which relied on “rhyme, stress alternation, and syllable counting”, were becoming widespread in ME (Minkova 2012: 121, cf. Minkova 2007: 183). Wittig (1978: 5), for example, shows that the metrical forms

used by ME romances range from the octosyllabic line to the tail rhyme to the alliterative line, while Minkova (2012: 121) argues,

[v]erses of equal numbers of syllables – ‘isosyllabic verses’ – are not uncommon in [OE] poetry, but the recurrence was not structurally regulated; a verse could have from a minimum of four to fourteen syllables. Isosyllabism is an imported metrical feature in [ME]. Schemes based on the iteration of isosyllabic lines – the octosyllabic line, the septenarius, and, with Chaucer, the decasyllabic iambic pentameter – are at the core of [ME] verse composition.

Schemes based on the repetition of syllables, referred to in this study as ‘isosyllabic verse’, had a number of properties in common: a) they permitted an unstressed syllable after the last ictus, i.e. metrical stress, but since such syllables were outwith the metrical scheme, their presence or absence did not affect the isosyllabicity of the line; b) the lines of isosyllabic verse were often grouped in couplets or larger groups of end-rhymes; and c) although rhyming sporadically occurred as an ornamentation in OE verse, the influence of Anglo-Norman made this stylistic device “the verse-line marker of choice” (Minkova 2012: 121). However, while the above features characterise all of the schemes mentioned by Minkova (2012) in the above quotation, the decasyllabic iambic pentameter, which was based on the iteration of isosyllabic lines, i.e. the iamb, differed from the other ones in that it was based on a binary sequence of a weak and strong position (W S). Although iambic feet could occur in OE, the iamb turned into the dominant metrical foot in ME, being featured in such works as *The Owl and the Nightingale* and *The Ormulum* (Minkova 2012: 121-122), in addition to *Genesis and Exodus*, *Handling Synne* by Robert Mannyng of Brunne, *Sir Orfeo*, *Havelok the Dane*, *Cursor Mundi*, Barbour’s *The Bruce* and Gower’s *Confessio Amantis*. The use of the iamb involved both stress (accent) and syllable, whereby an unstressed syllable, marked with ‘x’, alternates with a stressed one, marked as ‘/’, to comprise a foot, the sequence of which, in turn, forms a line. Other common feet in English beside the iamb (x /) included trochee (/ x), pyrrhic (x x) and spondee (/ /), with a sequence of four feet being a tetrameter and so on (J. Smith 2012: 56). Nevertheless, before “Chaucer started using five feet ..., [ME] isosyllabic poetry was written in eight-syllable [or octosyllabic lines] ... linked in couplets or larger groups by end rhymes” (Minkova 2007: 183).

Octosyllabic verse, also known as short-line, probably developed under the French influence, combining two traditional parameters: French position size (one syllable) and Germanic position number (eight). However, this kind of verse practised in ME was less regular, resulting in many lines having seven syllables, not eight, and a trochaic, rather than an iambic rhythm (Duffell 2008: 78). In relation to Scots poets versifying in the 14th century Scottish short-line (which might have been closest to NME verse, too), Duffell (2008: 97) argues that they were schooled in mainland French, rather than insular French. This had an effect on the shape of the octosyllabic line in that syllables were more regular, despite their similarity of verse to the southern contemporaries in terms of the number of beats or stresses. This better controlled regularity of syllables accounts for, for example, the “exceptionally regular” line in Barbour’s *The Bruce* being octosyllabic, with “94% per cent of all lines in the poem [having exactly eight syllables], a figure significantly higher than that found in any short-line poem by Chaucer” (Duffell 2008: 97). In a similar way, *The Scottish Legendary*, which includes *St. Andrew and the Three Questions* and *The Life of St. Julian Hospitaller*, both of which are used in this study as sources for the ESc data set, is written in “[t]he rhyming tetrameter (4-beat) couplets ... managed for the most part with fluency, apart from some clumsy inversions of natural word order” (Whatley, Thompson & Upchurch 2004a; 2004b). However, although the view adopted here is that the works produced in ESc were intended to be written with an equal number of syllables per line, the process of counting syllables necessitates references to phonetic processes, such as the loss of final <e> in NME and ESc (J. Smith 1996, Duffell 2008), but also elision and syncopation, as discussed in Minkova (2007: 184-186), which are not reflected by orthographic representation.

The octosyllabic line quickly established itself as the canonical short-line in Scots, with other types of line used in this variety, too, including the long-line, which was based on the alliterative verse similar to *Sir Gawain and the Green Knight*, but taking longer to emerge (Duffell 2008: 97-99) (Duffell 2008: 98-99, for example, argues that long-line, consisting of a strophe of thirteen lines subdividing 8 + 5, was employed in a number of English and Scottish alliterative poems of the 14th and the 15th centuries. Both strophes are constructed using the following principle: an octave

consists of lines with four measures, preceding a bob and wheel containing shorter lines with fewer measures. The rhyme scheme is ABABABAB CDDDC. For the metrical constituents of ME and Scots alliterative verse, cf. Minkova 2007: 177-181, Duffell 2008: 98-99, respectively). In terms of the type of the literature that used the octosyllabic line in general, Damian-Grint (1999: 186) argues, although referring to French romances, which were generally characterised by this kind of metre, at least initially, “[a]s a ‘general-purpose prosody’ ... the octosyllabic couplet was a form proper to ‘serious literature’ – religious, didactic and historical works – and its use in romances indicates an attempt to give them a similar seriousness and ‘truthfulness’”. However, in contrast to Scottish poets, most poets in England at the time favoured the rough octosyllable, characterised by some irregularities. K. Campbell (1975: lxxvi), for example, notes that MSS like Cotton and Rawlinson, which contain extracts from *Cursor Mundi*, “[exhibit] all the freedoms and irregularities of meter that were allowed in the better poetry of the time, and it also affects certain licenses which the better poet shunned”. Let us, therefore, investigate the occurrence of **gan** and **can** from the perspective of the type of line or metre in verse texts in the NME and ESc data sets.

The results of this study show that the lines containing **gan** and **can** in the NME and ESc data sets can be predominantly classed as written in the octosyllabic line in Period I, as in (14) below and (15) below, but, additionally, in trisyllable in Period II, as in (16) below. By comparison, the occurrence of **beginnen** in verse works in these data sets is found in the same types of line, but also in lines of other types. For example, unlike **gan** and **can**, this verb is found in *The Wakefield Pageant in the Towneley Cycle*, which is written in the Wakefield stanza¹⁰², as in (17) below. The octosyllabic line containing instances of **gan** and **can** is found in the following texts

¹⁰² The Wakefield stanza consists of nine lines in the pattern of AAAA/BBBB CDDDC (Stevens 1981: 101, Arnovick 2012: 570). The a-rhyme consists of the internal rhyme in a quatrain (frons), whereas the b-rhyme functions as an end-rhyme there. The c-rhyme forms a one-stress (or sometimes two-stress) bob, similar to that used in *Sir Gawain and the Green Knight*; this bob rhymes with the last line of the stanza, this sometimes inaccurately being referred to as tag. The d-rhyme, on the other hand, occurs in the middle three lines in the cauda section, the term referring to the five short lines ending the stanza (Stevens 1981: 101). Stevens (1981: 101) also observes that the cauda lines normally consist of two stresses, with a variable number of unstressed syllables, the same length as the half lines of the frons. However, this stanza is also interpreted as being thirteen lines in length by Stevens (1981) and Stevens & Cawley (1994) in the pattern of ABABA BAB CDDDC, containing a cross-rhymed octet frons, with a tercet cauda with tail-rhymes.

used in the NME and ESc data sets: *Cursor Mundi*, *Octavian*, *Sir Eglamour of Artois*, *St. Andrew and the Three Questions*, *The Life of St. Julian Hospitaller*, *The Northern Homily Cycle*, *The Pricke of Conscience* and *The York Plays*. However, there are some deviations from the type of line in some of these works. For example, *The York Plays* is characterised by strophes of different length and varying metre, typically consisting of a quatrain of tetrameter and a series of shorter lines (Duffell 2008: 139), although the strophe containing *gan* with infinitive in this work from the NME data set consists of a tercet of eight syllables, followed by a single tetrasyllabic line. *Sir Eglamour of Artois* and *Octavian*, on the other hand, have instances of *gan* and *can* in trisyllabic lines (found occurring in the tail). Therefore, despite the findings from the data sets suggesting that this verb and its variant predominantly occur in the octosyllabic line in the NME and ESc data sets, the presence of *gan* and *can* there could just reflect the main use of this type line at the time. Furthermore, the predominant presence of this verb and its variant in lines characterised by eight syllables could also be accidental, rather than having anything to do with the precise line or metre being used – further research would need to be carried out to confirm this.

- (14) *Pe*¹ *slogth*² *he*³ *fand*⁴ *pat*⁵ *him*⁶ *gan*⁷ *wiss*⁸
 The path he found that him did mark
*Til*¹ · *ward*² *pe*³ *zate*⁴ *of*⁵ *pa*⁶ · *ra*⁷ · *dis*⁸;
 Toward the gate of paradise;
 (1350-1420 *Cursor Mundi* p. 82; *HCET*)

‘He found the path that marked the way for him towards the gate of paradise;’

- (15) *Ful*¹ *wel*² *war*³ *pat*⁴ *wit*⁵ *him*⁶ *pat*⁷ *night*⁸
 Full well were they with him that night
*Pe*¹ *morn*² *his*³ *ass*⁴ *pan*⁵ *can*⁶ *he*⁷ *dight*⁸,
 The morn his ass then did he saddle,
 (1350-1420 *Cursor Mundi* pp. 1011-1012; *HCET*)

‘They were able to be with him that night. The next morning he saddled his donkey’

- (16) *Sche*¹ *wend*² *hyt*³ *had*⁴ *be*⁵ *byg*⁶ · *gyd*⁷ *londe*⁸
 She believed it had been inhabited land
*And*¹ *there*² *up*³ *gan*⁴ *sche*⁵ *wende*⁶

And there up did she go
 (1420-1500 *Sir Eglamour of Artois* l. 812-813; *TEAMS*)

‘She believed this had been an inhabited land and [so] she went there’

(17) (^ (\Vxor.\) ^)
 Then¹ be² · gyn³ · nys⁴ to⁵ grufe⁶ (gru⁶ · fe⁷) to⁷⁽⁸⁾ vs⁸⁽⁹⁾
 Then begins to grow to us
 me⁹⁽¹⁰⁾ · ry¹⁰⁽¹¹⁾ che¹¹⁽¹²⁾ · re¹²⁽¹³⁾.
 merry cheer.
 Bot¹, hus² · band³,
 But, husband,
 What¹ grownd² may³ this⁴ be⁵?
 What ground may this be?
 (1420-1500 *The Wakefield Pageants in the Towneley Cycle* p. 26;
HCET)

‘Then cheerfulness will soon begin for us. But, husband, what country may this be?’

While the lines containing **gan** and **can** are mainly octosyllabic in the NME and ESc data sets, the following two pieces written in verse in the NME data set deserve some attention in relation to the type of line or metre used in them: *The Wakefield Pageants in the Towneley Cycle* and *Meditatio de passione Cristi*. With reference to the former, Chambers (1947: 37) maintains that this text is “written in a characteristic metre of its own ... [t]he rhythm [being] markedly anapaestic”, although, as already mentioned above, the type of metre used in this work is not to suggest that this is the reason why neither **gan** nor **can** occurs there. In relation to the latter, on the other hand, the metre of *Meditatio de passione Cristi* appears to be more complex; namely, this piece appears to have strophes of different length, i.e. couplets, tercets and quatrains, describing different stages of the Passion of Christ, with half-lines being either five syllables (optionally six) or seven syllables (optionally eight) long (not counting the final <e>, being here orthographic, for example in *blode* in (19) below, which was being lost in unstressed words in the 11th century and disappeared in stressed words in the 13th century in the northern varieties of ME, and in the 14th century in Scots, as argued by J. Smith 1996: *Intralinguistic Correspondences I: The Loss of Final -E* and Duffell 2008: 96, respectively). However, the metre of the half-lines containing the instances of **gan** with infinitive (given the frequency of rhyme at the end of every half-line, as shown in (8) above

and (9) above) can, to an extent (i.e. taking into account the above variation), be interpreted as consisting of three iambs, as in (18) below, but with some irregularities in (19) below, the difference being six syllables in the former, and seven and five syllables in the latter. The reason for this irregularity is that phonological features of ME made accentual verse more suitable for the language than any stress-syllable metre, especially when considering the fact that accentual metre is indifferent to the elision of the final <e> (the term ‘elision’ refers to situations when “[t]he final vowel of an unstressed syllable can be elided when adjacent to another vowel, both across a word-boundary and within a word” Minkova 2007: 184).

(18) x / x / x /
Ful fast | þai gan | hym dyng,
 Full fast they did/began to him strike,
 x / x / x /
and at | þe py · | ler swyng,
 and at the pillar swing,
 (1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 57;
ICoMEP)

‘They quickly did/began to strike him and swung him at the pillar,’

(19) x / x / x / x
In fyue | sted · es | of his flesch
 In five steads of his flesh
 x / x / x
þe blode | gan downe | glyde,
 the blood began to down glide,
 (1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 57;
ICoMEP)

‘In five places of his body, the blood began to glide down,’

In the literature on the subject, it is possible to come across various terminologies when referring to isosyllabic verse, such as the octosyllabic line (for the structure of ME alliterative verse, cf. Minkova 2007: 177-184; for the principles of OE alliterative verse, cf. Minkova 2012: 116-117). For example, Minkova (2007: 186) refers to “[t]he [ME] iambic *tetrameter*, or the four-beat/octosyllabic line, as it is alternatively known”, but she also provides an explanation: the term ‘four-beat’ bears resemblance to the four-beat stress of OE verse and refers to the practice of

alliterative verse in ME; the term ‘octosyllabic’, on the other hand, stresses the idea of uniformity of the number of syllables per line, whereas the term ‘iambic tetrameter’ refers to the number of beats as a notion of measure (Minkova 2007: 186-187). Therefore, the use of such terminology depends on the classification of the line of verse into different units at different levels. According to Minkova (2007: 187), the division of the line was into two parts known as colons, which, in turn, subdivided into two feet, each consisting of syllables. Nevertheless, the division into metrical feet can also mean that those at the left and right edge of the line are characterised by different properties – the first foot may lack a filled weak position, rendering it defective and making the line ‘headless’. A reversal of the first foot, i.e. strong and weak instead of weak and strong can occur, while this kind of reversal is prohibited in the last foot of the verse (Minkova 2007: 187). In view of such a reversal, it is, therefore, crucial to remember the difference between metre and rhythm. In the rhythm of speech, stress patterns are assigned in accordance with the category of the item in question. J. Smith (2012: 56), for example, demonstrates that in such cases, words that perform a grammatical function are less stressed than those belonging to the open class (lexical). Strachan & Terry ([2000] 2011: 72) summarise the distinction between metre and rhythm thus,

[i]t is helpful to think of metre as a subset of the larger concept of rhythm and, consequently, to consider formal rhythmical analysis of recurring poetic sound-patterning as indistinguishable from metrical analysis. Thus metre is a specific type of rhythm, and might best be defined as the ‘measurable sound-pattern evident, in varying degrees of regularity, in a line of poetry.

Some lines containing instances of *gan* and *can* in the NME and ESc data sets also require a reversal of the first foot. For example, although the iamb (x /) should be used as a metrical foot in (20) below, this example constitutes an instance of the first foot lacking a filled weak position. As a result, this initial foot looks like a trochee (/ x), but with a reversal, it should be realised as an iamb to suit the metre (this reversal is marked with a twisted line in the example below).

(20) / x x / x / x /
 A · dam, | thy · selffe | made al | þis syte,
 Adam, thyself made all this grief,
 / x x / x / x /
 For to | the tree | þou wente | full tyte
 For to the tree thou went full quickly
 x / x / x / x /
 And bolde · | ly on | the frute | gan byte
 And boldly on the fruit did/began to bite
 x / x /
 My lord | for · bed
 My lord forbade
 (1420-1500 *The York Plays* p. 70; *HCET*)

‘Adam, you made all this grief yourself, because you very quickly went to the tree and boldly did/began to bite on the fruit that my lord forbade’

6.2.3 Alliteration and Rhyme

Verse is also characterised by sound patterns, such as alliteration and rhyme.

However, *gan* and *can* occur infrequently in unrhymed, alliterative verse in ME and Scots (Funke 1922: 22, Koziol 1932: 132-133, Mustanoja 1960: 612, Visser 1969: 1572, Terasawa 1974: 99-100, Tajima 1975: 434, 437, Brinton 1983: 236; 1988: 120-121). In OE and in ME and Scots, alliteration had different functions. In OE, the pure-stress poetry correlated with the structure of the language at the time; determiners were largely optional, auxiliary verbs had an uncertain status, whereas the function of inflectional endings was to join units together. Given the dominant metrical foot in OE was the trochee (/ x), as opposed to the iamb (x /), which became more popular in ME, alliteration had a function of providing the linking between metrical units (feet), i.e. it was a cohesive device, in addition to the stress-pattern (J. Smith 2012: 58). However, although alliteration was becoming to be used for expressive or ornamental purposes in the rhyming verse in ME, as discussed below, in alliterative verse, there is still evidence of structure. For example, J. Smith (2012: 58-59) argues that in this kind of verse, “the primary poetic signifier was alliteration, whereby alliterative lines consisted of the pattern *aa//ax* as the norm, with deviation, especially but not always in the first half-line (the ‘A-stave’), offering the opportunity of variation” (for a more detailed discussion of alliterative patterns of

ME verse, cf. Minkova 2007).¹⁰³ He provides an example of this kind of structural alliteration from Dunbar’s narrative poem *Ane Tretis of the Twa Mariit Wemen and the Wedo*, as in (21) below, whereby the first three stressed words alliterate, i.e. *bird*, *bransche* and *birst*, with the last stressed word, *notis*, not, all following the pattern *aa//ax*.

- (21) *Quahairon* *ane* *bird* *on* *ane* *bransche* *so*
 Whereon a bird on a branch so
birst *out* *hir* *notis*
 burst out her notes
 (Dunbar *Ane Tretis of the Twa Mariit Wemen and the Wedo* l. 5; as
 cited in J. Smith 2012: 59)

‘And then, a bird on a branch burst out its musical notes’

Although the practice of alliteration found in OE poetry continues to an extent in ME and Scots in the form of alliterative verse, it is also possible to observe some changes with regards to the use of this stylistic device in other types of verse, more specifically rhyming verse typical of the ME and Scots poetry of the time. In rhyming verse, as opposed to alliterative verse, a tendency for alliteration was to be increasingly used for expressive or ornamental, rather than structural, purposes (cf. Blake 1977, Pearsall 1977: 84, Minkova 2003: 15). Oakden (1930), for example, provides evidence from a range of eME texts to support his view that the function of alliteration was increasingly used “to make its appeal to the eye as well as to the ear” (178) – in contrast to OE, unstressed pronouns, auxiliary verbs, the second element in a compound, verbal prefixes were found in alliteration in various eME and ME texts (the poems of the Alliterative Revival) with a varying frequency of occurrence. Oakden (1930: 179) also notes another tendency for alliteration already found in eME: “the tendency to excessive alliteration”, about which he comments,

[t]his increase in the actual amount of alliteration used is indicative of the direction which the alliterative poetry is taking. Alliteration was destined to become more and more ornamental, and less and less structural. Incidentally this is part of the larger tendency we have already discussed [in the above work], for the alliteration to make its appeal to the eye as well as to the ear. In *Sir Gawain and the Green Knight*, for example, we find many lines with super-abundant alliteration of a purely ornamental kind. (Oakden 1930: 179)

¹⁰³ In this pattern, *aa* stands for the alliterating stressed words, whereas *x* denotes the non-alliterating stress word; // signifies a medial caesura (J. Smith 2012: 59).

According to Oakden's (1930: 236) data, the evidence of ornamental alliteration appears to be strongest in texts written in the Midland dialect. Although he lists *The Turke and Gawain* (c 1450) from the "Nth." dialectal area as containing "alliteration in varying degrees", he also points out a disproportionate use of ornamental alliteration amongst a mid-14th century school of poets in the northern counties and in Scotland, flourishing for two centuries,

[t]hese writers were evidently accustomed to the alliterative long line with end-rhyme, as used in lyrical verse. They desired, however, to write longer poems and possibly recognised the unsuitability of the poetic medium at their disposal. The simple stanza of eight or more metrically uniform lines did not prove very attractive, and this difficulty was obviated by the use of the *bob and wheel* or of the *wheel* alone. The resulting stanza-forms became very popular in the north, but never in the west and elsewhere. The experiment was not unsuccessful, yet looking back we are able to see the danger involved. It was the desire for ornament that led these poets to overcrowd their lines with alliterating sounds, and to employ so many metrical devices. In its last stages the long line is a mere jingle of sounds. Possibly the northern and western schools of alliterative poets were rivals. (Oakden 1930: 244-245)

Evidence of alliteration used for expressive or ornamental reasons can be found in Barbour's *The Bruce*. Eyre-Todd ([1907] 1996: v) argues that other ESc poems, namely *The Taill of Rauf Coilzear*, *The Awntyrs of Arthure*, *Sir Tristrem* and *The Pystyl of Swete Susan*, are written in "the more inflected language and in the alliterative and accented verse-forms of an earlier time." *The Bruce*, on the other hand, "definitely committed the poetry of Scotland to metre and rhyme, instead of the older alliteration and accents, as its distinguishing features." To support the role of alliteration being ornamental in *The Bruce*, references can be made to Purdie (2015: 59-60, 65), who shows that alliteration is not consistently used in this work, with this literary stylistic device being very frequent in some passages in this work, perhaps being an inspiration for Barbour from OF *romans antiques*¹⁰⁴. Given this

¹⁰⁴ Purdie (2015: 60) argues that the earliest versions of *romans antiques* dealt with important ancient historical events, reinforcing the status of such events "as quasi-historiographical works by basing – or claiming to base – themselves on authoritative Latin histories". Such versions of *romans antiques* would also contain some elements of the romance genre,

the emphasis on illustrating courtly manners alongside chivalric prowess; the romantic love stories, often absent from their sources, slipped into the wider 'historical' narrative ... These *romans* were also vast, diffuse in focus (all offer a selection of heroes' careers to follow),

lack of consistency of alliteration in *The Bruce*, Oakden (1930: 245) summarises the role of this stylistic literary device during the ME period, “[a]lliteration, once structural, had become more and more ornamental, until finally the verse became chaotic and ceased to be a worthy vehicle of what little poetic inspiration survived.” Overall, therefore, the rhyming verse, such as narrative verse in the octosyllabic line, would be characterised by alliteration for expressive or ornamental purposes.

As the sources listed at the start of this section argue that **gan** and **can** are infrequent in unrhymed, alliterative verse, my findings demonstrate that the texts where this verb and its variant occur in the NME and ESc data sets also show no evidence of alliteration used for structural purposes.¹⁰⁵ Instead, alliteration in such texts points to this device being used for expressive or ornamental purposes. So, in relation to the lines with **gan** in the NME data set, alliteration is found in *Meditatio de passione Cristi* in *Richard Rolle of Hampole (Yorkshire Writers)*, as in (22) below (given the problem of the arrangement of the lines of verse in this piece, I have marked alliteration in the whole example: /f/ in *five, flesch*; /g/ in *gan, glide*; and /s/ in *steads, stremes, strande*), but also to a varying degree in *Sir Eglamour of Artois*, as in (23) below (/s/ in *sqwyer* and *say*, although there is also /w/ in *walkest* and *west* in a line where there is no **gan**) and *Octavian*, as in (24) below (/ð/ in *thedir, than*, but also /a/ in *alle, appon* and *a*, and /ð/ in *the* and *that* in lines not containing **gan**), both texts containing high frequencies of occurrence of **gan** (cf. Section 2.4.1.1.1). Other texts from the NME data set containing the verb **gan** and alliteration include *Cursor Mundi*, *The Northern Homily Cycle*, *The Pricke of Conscience* and *The York Plays*. With reference to **can**, alliteration is present in the same line as this variant in *Cursor Mundi* in the NME data set, as in (25) below (/k/ in *cuthli, can* and *knele*, but also /w/ in *was, wist* and *wel* elsewhere in the example) and in *St. Andrew and the Three Questions* in the ESc data set, as in (26) below (/ð/ in *that* and *this*, but also /a/ in *And*

interested in political contexts and consequences, and finally, proud of the fact that they were dealing with allegedly historical material. They also featured the kinds of tragic ending that are often now considered to be anathema to the genre of romance ... These *romans antiques* were pioneers of what became the new genre of romance[.] (Purdie 2015: 60-61)

¹⁰⁵ With no instances of **gan** and **can**, *The Wakefield Pageants in the Towneley Cycle* does not use alliteration in a systematic way: “very little alliteration ... The great mass of the stanza is ... without it. In others its use is seemingly accidental, as in the first and second stanzas quoted. But sometimes, when a serious or solemn effect was desired, its use appears to have been intentional” (Cady 1912: 247).

and *at* in another line). While the repetition of initial consonants is found on items in the lines containing *gan* and *can* (and beyond) in the NME and ESc data sets, alliteration can also be seen incorporating both this verb and its variant, as in (22) below and (25) below, but in isolated instances and to a lesser extent. Overall, therefore, both *gan* and *can* are found in verse works characterised by alliteration in the NME and ESc data sets, but this stylistic device is used in such texts for expressive and ornamental, rather than structural, purposes.

- (22) *In fīve steads of his flesh þe blode gan
 In five steads of his flesh the blood began
 downe glyde, als streμες of þe strande
 down glide, as streams of the river
 (1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 57;
ICoMEP)*

‘In five parts of his body, the blood began to glide down, like streams of the river’

- (23) *The knyght to his sqwyer gan say
 The knight to his squire did say
 In chambur ther he gan reste,
 In chamber there he did rest,
 “Belamy, and thou kowdest leyne,
 “Friend, and thou could remain silent
 A counsell I wold to the sayne –
 A counsel I would to thee say –
 Thou walkest both est and west.”
 Thou walk both east and west.”
 (1420-1500 *Sir Eglamour of Artois* l. 50-54; *TEAMS*)*

‘The knight said to his squire in [the] chamber, where he rested [was resting], ‘Friend, if you could keep a secret, I offer advice to you – you travel widely.’

- (24) *The dir than gan he bede.
 Tither then did he summon.
Alle were thay sampnede appon a daye
 All were they assembled upon a day
 With grete solace and mekill playe;
 With great comfort and much pleasure;
 To the kyrke that lady yede.
 To the kirk that lady went.
 (1420-1500 *Octavian* l. 190-195; *TEAMS*)*

‘He then summoned to that place. They were all assembled with great comfort and much pleasure within the space of a day; that lady went to the church.’

- (25) *Pat it was iewus þan wist i wel,*
 That it was Jesus then knew I well,
And cuthli for him can i knele,
 And certainly for him did/began to I kneel,
And i said, ‘ai lauerd! blisced þou be
 And I said, ‘Aye Lord! Blessed you be
 (1350-1420 *Cursor Mundi* p. 1014; *HCET*)

‘I knew well that it was Jesus, and certainly I did/began to genuflect/kneel before him, saying, ‘Yes, Lord! Blessed be you’

- (26) *Sad to hyme agane, “Thu ga*
 Said to him again, “Thou go
Til hyre that cane this demand ma
 To hir that did this demand make
And spere at hyre grathly.
 And ask at her promptly.
 (1375-1420 *St. Andrew and the Three Questions* l. 1084-1087;
TEAMS)

‘[the porter] said to him again, “Go to her that made this demand and ask her immediately.”’

As a sound-patterning device, rhyme aurally marks the ends of lines in verse to create a structural link between lines. The occurrence of *gan* and *can* in rhymed verse appears to be significant for our understanding of the role of this verb and its variant in ME and Scots poetry – one of the functions of this verb and its variant is to enable the verb in infinitive to provide rhyme in a line of verse (Koziol 1932: 132-133, Kerkhof 1966: 154, 158, Smyser 1967: 74, Terasawa 1974: 102, Tajima 1975: 434, 437-438, Brinton 1981: 164). However, an overview of the literature on the occurrence of *gan* and *can* in ME and Scots verse does not suggest any particular type of rhyming scheme being associated with this verb and its variant. Nonetheless, with this verb and its variant being typically found in narrative verse texts, as discussed in Section 6.2.1, Terasawa (1974: 100) finds higher frequencies of *gan* and *can* in rhymed ME and Scots romances written in “short couplets”, a higher average

in “tale-rime”¹⁰⁶ romances, with “[o]ther stanzaic romances ... also inclin[ing] to favour the *gan* idiom”, while Ward, Waller, Trent, Erskine, Sherman & van Doren ([1907-21] 2000: §8) list “rhyme royal”¹⁰⁷ for *Troilus and Criseyde*, in which Chaucer uses *gan* “more frequently than in any other text” (Brinton 1996: 75). Koziol (1932: 132-133), Brinton (1988: 121) and Sims (2008: 140) have “bob and wheel”¹⁰⁸ as the type of verse in which *gan* is found, too, although it is clear that this type of construction is not restricted to this type of line throughout, especially in texts like *Sir Gawain and the Green Knight*. Overall, therefore, the use of *gan* and *can* in stanzas of a specific kind appears to vary according to the type of text using this kind of stanza.

In the NME and ESc data sets, between 88% to 100% of instances of the verb in infinitive with *gan* and *can* occur at the end of a line, as part of rhyme, as outlined in Table 54 and Table 55, respectively (the observed relative frequencies are per the number of words in verse texts only, cf. Section 2.2.3.2).

¹⁰⁶ The term ‘tale-rhyme’ refers to a stanza frequently used in romances. The most common types of stanza found are the six-line stanza AABCCB and the twelve-line form AABCCBDDBEEB initially found in *Amis and Amiloun* (Putter 2009).

¹⁰⁷ The term ‘rhyme royal’ (or ‘rime royal’) refers to a stanza consisting of seven lines written in iambic pentameter. The rhyme pattern is ABABBCC. This type of rhyme became to be used in long narrative poems during the 15th and the 16th centuries. However, it was first used by Chaucer in 14th century in *Troilus and Criseyde* and *The Parliament of Fowls* (*Merriam-Webster’s Encyclopedia of Literature* s.v. *rhyme royal*).

¹⁰⁸ The term ‘bob and wheel’ refers to the following: a five-line rhyming section is added at the end of unrhymed lines ranging from twelve to thirty seven; this rhyming section consists of one syllable bob, followed by a wheel of four lines of three syllables, the whole rhyming ABABA (Turville-Petre 1977: 62).

Table 54

*Position of the Verb in Infinitive with **gan** and **can** in a Verse Line in the NME Data Set in Period I and Period II*¹⁰⁹

	<i>gan</i>						<i>can</i>					
	Verb in Inf						Verb in Inf					
	Rhyme			Non-rhyme			Rhyme			Non-rhyme		
	<i>N</i>	<i>R</i> ₃	%	<i>N</i>	<i>R</i> ₃	%	<i>N</i>	<i>R</i> ₃	%	<i>N</i>	<i>R</i> ₃	%
Period I	17	69	100	-	-	-	7	29	88	1	4	12
Period II	73	244	97	2	7	3	12	40	100	-	-	-

Notes: *N* – Observed absolute frequency
*R*₃ – Observed relative frequency per verse sample per 100,000 words
 % – Per-cent of verbs in infinitive in rhyme or non-rhyme position
 NB: Observed relative frequency is rounded to the nearest integer

Table 55

*Position of the Verb in Infinitive with **can** in a Verse Line in the ESc Data Set in Period I*

	<i>can</i>					
	Verb in Inf					
	Rhyme			Non-rhyme		
	<i>N</i>	<i>R</i> ₃	%	<i>N</i>	<i>R</i> ₃	%
Period I	8	216	100	-	-	-

Notes: *N* – Observed absolute frequency
*R*₃ – Observed relative frequency per verse sample per 100,000 words
 % – Per-cent of verbs in infinitive in rhyme or non-rhyme position
 NB: Observed relative frequency is rounded to the nearest integer

However, while the frequency represented by the number of instances of the verb in infinitive with **gan** appears to be subject to a decrease from that denoting this verb in rhyme position to that in non-rhyme position in Period II (the opposite can be seen

¹⁰⁹ This includes the two instances of **gan** with the plain infinitive in *Meditatio de passione Cristi*.

for the verb in infinitive with *can* in Period I), this effectively concerns isolated instances in verse texts in the NME data set. Nevertheless, this decrease is statistically significant, while the increase in the use of the verb in infinitive in non-rhyme position, when found with *gan*, is not, as discussed in more detail below. In relation to *can*, on the other hand, there is an increase in the application of the verb in infinitive in rhyme position over the periods in verse texts in the NME data set, with a decrease in the use of the verb in infinitive in non-rhyme position. However, this increase in the verb in infinitive in rhyme position and decrease of the verb in infinitive in non-rhyme position has no statistical significance, as also detailed below. (It should be observed that while the verb in infinitive with *gan* and *can* is almost exclusively found in end rhyme position in verse texts in the NME and ESc data sets, the two instances of *gan* in Rolle's *Meditatio de passione Cristi* can form either an internal rhyme and end rhyme, as in (27) below, or an end rhyme, as in (28) below, although, effectively, either still in rhyme position¹¹⁰).

- (27) *Ful fast þai gan hym dyng, and at þe pyler swyng,
& his fayre face defowlyng / with spittyng. /
(1350-1420 Richard Rolle of Hampole, Yorkshire Writers p. 57;
ICoMEP)*

'They quickly did/began to strike him and swung him at the pillar, and defiled his beautiful face with spitting.'

- (28) *Ful fast þai gan hym dyng,
Full fast they did/began to him strike,
and at þe pyler swyng,
and at the pillar swing,
& his fayre face defowlyng /*

¹¹⁰ To find some literary devices like rhyme in Rolle's English prose works in general is not unusual, though. Allen (1989: 36) argues that "even in his prose [Rolle] constantly has recourse to the rhetorical device of *similiter desinens*, words ending with an identical grammatical inflection, or *similiter cadens*, where the sound is the same but the grammatical function is different." Rolle applied these devices to his English, not just Latin, works. What is more, although not referring to any of Rolle's works in particular, Allen (1989: 37) points out, "[y]et such devices are less frequent in the English work: the rhetoric and rhymes are simpler, and the tone of the lyrics more petitionary and less ecstatic than in the Latin." Quoting Riehle (1981: 5), Allen (1989: 37) argues that the reason for the less frequent uses of such devices in Rolle's works is his confidence in formal Latin, rather than English, training. However, she also adds that "the simpler diction of the English works is a deliberate ploy of Rolle's", as his English writings are aimed at "'untaught,' women, for the most part, who do not know Latin, who might be confused by verbal artistry if it were too elaborate or frequent, and who cannot even easily read English if the words are too long or the vocabulary too Latinate" (Allen 1989: 37).

& his fair face defouling /
with spittyng. /
with spitting. /
(1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 57;
ICoMEP)

‘They quickly did/began to strike him and swung him at the pillar, and defiled his beautiful face with spitting.’

Statistically, the null hypothesis is that there is no relationship between the differences in the frequency of occurrence of the verb in infinitive of **gan** and **can** in either rhyme or non-rhyme position over Period I and Period II in verse texts in the NME and ESc data sets. In relation to the position of the verb in infinitive with **gan** in a line of verse in the NME data set, the log-likelihood value is 27.31 for the verb in infinitive in rhyme position, this result being statistically significant at $p < 0.0001$ (99.99th percentile, 0.01% level), showing that we should reject the null hypothesis in this case. In contrast, the log-likelihood value is 2.40 for this verb in non-rhyme position, suggesting that we should accept the null hypothesis at the level of $p < 0.05$ (95th percentile, 5% level). Conducting a statistical test to determine whether there is a correlation between the differences in occurrence of the verb in infinitive in either rhyme or non-rhyme position when found with **can** in verse texts in the NME data set, we can be sure that with the log-likelihood values being 0.53 and 1.59 (95th percentile, 5% level), we can accept the null hypothesis that there is no relationship between the differences in occurrence of the verb in infinitive in either rhyme or non-rhyme position over Period I and Period II at $p < 0.05$.

On the basis of the information in Table 54 and Table 55, it is clear that the verb in infinitive found with **gan** and **can** in the NME and ESc data sets prefers rhyme position in a line of verse. In terms of rhyme patterns, the following types of stanza can be identified in the verse texts where this verb and its variant are recorded in these data sets: a) an octosyllabic couplet in *Cursor Mundi*, *The Life of St. Julian Hospitaller*, *The Northern Homily Cycle*, *The Pricke of Conscience* and *St. Andrew and the Three Questions*, as in (29) below; b) a quatrain consisting of an octosyllabic tercet (where the verb in infinitive is found) and a line of dimeter in *The York Plays*, as in (30) below; and c) tail-rhyme, comprising a tercet, where there is an

octosyllabic couplet and a trisyllabic tail in *Sir Eglamour of Artois* and *Octavian*, with the verb in infinitive with *gan* and *can* being part of the couplet, as in (31) below, but also the tail, as in (32) below (as argued by Pearsall 2011: 9, this type of stanza was almost exclusively used for romances from the middle of the 14th century, replacing the French-derived octosyllabic couplet of earlier ME narratives). By far, the most popular type of rhyme in which the verb in infinitive provides rhyme appears to be the couplet, with 63% of the texts in which *gan* and *can* are found belonging to this group. However, this percentage value excludes the texts which are characterised by more complex rhyme patterns like tail-rhyme, but which can still contain instances of the verb in infinitive in the couplet, as in (31) below.

- (29) *Sad to hyme agane, "Thu ga*
Said to him again, "Thou go
Til hyre that cane this demand ma
To hir that did this demand make
And spere at hyre grathly.
And ask at her promptly.
 (1375-1420 *St. Andrew and the Three Questions* l. 1084-1087;
TEAMS)

‘[the porter] said to him again, “Go to her that made this demand and ask her immediately.”’

- (30) *Adam, thyselffe made al bis syte,*
Adam, thyself made all this grief,
For to the tree þou wente full tyte
For to the tree thou went full quickly
And boldelyon the frute gan byte
And boldly on the fruit did/began to bite
My lord forbed
My lord forbade
 (1420-1500 *The York Plays* p. 70; *HCET*)

‘Adam, you made all this grief yourself, because you very quickly went to the tree and boldly did/began to bite on the fruit that my lord forbade’

- (31) *And the walles downe gan he dynge.*
And the walls down did/began to he strike.
And than gane alle the pepille crye.
And then did/began to all the people call.
Unto God and to mylde Marye
Unto God and to mild Mary

With sorowe and grete wepynge.
 With sorrow and great weeping.
 (1420-1500 *Octavian* l. 883-886; *TEAMS*)

‘And he did/began to break the walls down and then all the people did/began to call to God and to merciful Mary with sorrow and great weeping.’

(32) *The dir than gan he bede*
 Tither then did he summon.
Alle were thay sampnede appon a daye
 All were they assembled upon a day
With grete solace and mekill playe;
 With great comfort and much pleasure;
To the kyrke that lady yede.
 To the kirk that lady went.
 (1420-1500 *Octavian* l. 190-195; *TEAMS*)

‘He then summoned to that place. They were all assembled with great comfort and much pleasure within the space of a day; that lady went to the church.’

As also shown in Table 54 and already referred to above, there are three examples of the verb in infinitive found with *gan* and *can* not occurring in rhyme position in verse texts in the NME data set. I provide here more detail on these instances, adopting a view that it is possible to speak about the verb in infinitive found with *gan* and *can* having a tendency, rather than a strict rule, to occur at the end of a verse line. The occurrence of the verb in infinitive in non-rhyme position, therefore, involves the presence of another constituent, either modifying or complementing the verb in infinitive: the NP *pite*, which complements the verb *haf* in infinitive, as in (33) below, the NP *his chere* ‘his mood’, as in (34) below, and the AdvP *wele sore*, as in (35) below. Both *pite* and *wele sore* rhyme with lexical items found at the immediately preceding or following lines, but there are some points to note about their rhyme. In relation to *pite*, the final <e> of this NP was probably pronounced here to supply rhyme to the verb *be* in the preceding line, even though J. Smith (1996: Intralinguistic Correspondences I: The Loss of Final -E) argues that the final <e> was being lost in the unstressed words from the 11th century onwards and disappeared in stressed words in the 13th century in the northern varieties of ME. Furthermore, this NP was probably stressed on the final syllable at first, reflecting

the French stress pattern, only later adopting ‘native’, i.e. initial, stress. In relation to *sore*, on the other hand, both *sore* and *bare* do not rhyme on the final vowel, with the final <e> almost certainly not pronounced. Instead, we are presumably dealing here with scribal variation in the representation of OE long /a:/ (notice that *sore* and *bare* – the rounded southern variant in <sore> and unrounded northern variant in <bare> – could be orthographic variation, rather than reflecting actual pronunciation and/or be the result of copying by scribes from MSS from different dialectal areas in the history of the MS transmission). The NP *his chere* rhymes with *dere*, but which can be found four lines further down in the stanza containing these.

- (33) *Quen merci sagh him suagat be*
 When Mercy saw him thus be
Of him sco can haf pite,
 Of him she did/began to have pity,
Sco moght hir forbere nathing
 She might her[self] restrain nothing [not at all]
 (1350-1420 *Cursor Mundi* p. 550; *HCET*)

‘When Mercy saw him be in this way, she did/began to have pity, she could not restrain herself in any way’

- (34) *Than Clement appon the walles stode*
 Then Clement upon the walls stood
And full blythe he wex than in his mode
 And full happy he grow then in his spirit
And gan amend his chere
 And did amend his mood
And said, “Son, I hafe herde, I wene,
 And said, “Son, I have heard, I think,
 (1420-1500 *Octavian* l. 1007-1010; *TEAMS*)

‘Then Clement stood upon the walls and became happy in his soul; he cheered up, saying, “Son, I have heard, I think,’

- (35) *Clement so sorye was that daye*
 Clement so distressed was that day
For alle thaire costes that he solde paye,
 For all their gifts that he should compensate,
That he gane wepe wele sore.
 That he did/began to weep well sore.
And whills the kynges dauwnsede in the halle
 And whiles the kings danced in the hall

Clement tuke thaire mantills alle
 Clement took their cloaks all
 And to his howse tham bare.
 And to his house them bore.
 (1420-1500 *Octavian* l. 1208-1213; *TEAMS*)

‘Clement was so distressed that day, because he should compensate all their gifts, so he did/began to weep intensely. And while the kings danced in the hall, Clement took all their cloaks and carried them to his house.’

While it is clear from the findings presented here that the verb in infinitive found with *gan* and *can* almost exclusively occurs in rhyme position in verse texts in the NME and ESc data sets, the same cannot be said about the position of the verb in infinitive found with *beginnen* in verse texts in these data sets, as shown for comparison in Table 56 and Table 57 (the observed relative frequencies are per the number of words in verse texts only, cf. Section 2.2.3.2).

Table 56

*Position of the Verb in Infinitive with **beginnen** in a Verse Line in the NME Data Set in Period I and Period II*

	<i>beginnen</i>					
	Verb in Inf					
	Rhyme			Non-rhyme		
	<i>N</i>	<i>R₃</i>	%	<i>N</i>	<i>R₃</i>	%
Period I	6	25	50	6	25	50
Period II	13	44	81	3	10	19

Notes: *N* – Observed absolute frequency
R₃ – Observed relative frequency per verse sample per 100,000 words
 % – Per-cent of verbs in infinitive in rhyme or non-rhyme position
 NB: Observed relative frequency is rounded to the nearest integer

Table 57

*Position of the Verb in Infinitive with **beginnen** in a Verse Line in the ESc Data Set in Period I*

	<i>beginnen</i>					
	Verb in Inf					
	Rhyme			Non-rhyme		
	<i>N</i>	<i>R</i> ₃	%	<i>N</i>	<i>R</i> ₃	%
Period I	1	27	100	-	-	-

Notes: *N* – Observed absolute frequency
*R*₃ – Observed relative frequency per verse sample per 100,000 words
 % – Per-cent of verbs in infinitive in rhyme or non-rhyme position
 NB: Observed relative frequency is rounded to the nearest integer

The following observations can be made about the distribution of the verb in infinitive found with *beginnen* in verse texts in the NME data set: it occurs evenly in both rhyme and non-rhyme position in Period I; however, there is an increase in this verb being found in rhyme position at the cost of non-rhyme position in Period II, with a high percentage (81%) of the verb in infinitive in rhyme position in *Sir Eglamour of Artois* and *Octavian* in this period, with one instance being in non-rhyme position in *Octavian*, as in (36) below (where the verb *beginnen* itself occurs in rhyme position), and the remaining two instances being in *The Wakefield Pageants Towneley Cycle*. With the null hypothesis being that there is no correlation between the differences in the frequency of occurrence of the verb in infinitive in either rhyme or non-rhyme position in verse texts over periods in the NME data set, the log-likelihood values are 1.44 for the verb in infinitive in rhyme position and 1.70 for this verb in non-rhyme position – therefore, we can conclude there is no statistical difference in the frequency of occurrence of this verb in either position ($p < 0.05$, 95th percentile, 5% level) and so, we should accept the null hypothesis. In verse texts in the ESc data set, in contrast, only one instance of the verb in infinitive found with *beginnen* is recorded in rhyme position, with none in non-rhyme position.

(36) *Hys fadur was the furste man,*
 His father was the first man,

That he of bondys to lowse began
 That he of bonds to loosen/free did/began
 (1420-1500 *Octavian* l. 1747-1748; *TEAMS*)

‘His father was the first person¹¹¹, who he did/began to loosen/free of bonds’

In summary, *gan* and *can* is found in rhymed verse in the NME and ESc data sets, regardless of whether characterised by alliteration or not. Furthermore, although the verb in infinitive found with *gan* and *can* usually occurs in rhyme position, the verb *beginnen* shows no tendencies in general of making the verb in infinitive enable to occur in rhyme position in the vast majority of cases in these data sets, except in texts like *Sir Eglamour of Artois* and *Octavian*.

6.3 Proposals on Function(s) of *Gan* and *Can* in Middle English and Scots Verse Texts

Given the fact that *gan* and *can* occur in a non-ingressive sense and are mostly found with the plain infinitive, in addition to predominantly occurring in narrative verse, where the verb in infinitive found with this verb and its variant occurs at the end of a line to provide rhyme, a number of functions have been proposed in the literature for *gan* and *can* in ME and Scots, with varying degrees of acceptance and validity.

These functions are: a) metrical and rhyming; b) intensive-descriptive or stylistic; and c) narrative or textual. In the subsequent sections, I review claims in support of these functions. I also argue that the function of *gan* and *can* in the NME and ESc data sets is to enable the verb in infinitive to provide rhyme at the end of a line in a narrative verse text, in addition to this verb and its variant having textual and interpersonal functions in these data sets. Let us now turn to the most frequently cited role of *gan* and *can* in the literature, namely that of a metrical device.

6.3.1 Metrical and Rhyming Function

Given the occurrence of *gan* and *can* in verse in ME and Scots, it has been concluded that this verb and its variant must have a metrical function to provide an extra syllable in a line of verse (the function of this verb and its variant to enable the

¹¹¹ Given the fact that it is sometimes impossible to tell whether ‘man’ is intended to mean ‘person’ or ‘male human being’ (*OED* s.v. *man* n.¹), I have used the former as a more general term.

infinitive to provide rhyme is discussed below) (Taylor 1917: 574, Beschorner 1920: 15-19, Funke 1922: 22, Koziol 1932: 132-133, Kerkhof 1966: 154, 158, Smyser 1967: 68-69, Visser 1969: 1572, Terasawa 1974: 102, Tajima 1975: 434, 437-438, Mustanoja 1983, Fischer 1992: 266, Richardson 1991a: 80-81, Ogura 1998: 300, Sims 2008: 140, Fulk 2012: 106-107, cf. Brinton 1996). In this metrical function, *gan* and *can* occur as ‘a colourless filler’ (Funke 1922: 15; translation mine), i.e. as a meaningless device with no meaning. Statistical evidence appears to support this function: Smyser (1967: 68-69), for example, argues, “[o]ne thing may be made clear at the outset ... it is definitely a poetic form, not a prose one. In Chaucer’s verse, the various forms of [*gan*] appear nearly seven hundred times; in the very considerable bulk of his prose, they appear only three times”. However, the fact that *gan* is found in verse alone does not necessarily mean that the function of this verb is to provide an extra syllable, especially that sources identify other functions for this verb in verse, as discussed in this chapter. Furthermore, although Minkova (2007: 184) maintains that other constructions were ‘fillers’ in ME in that they provided an extra syllable in verse, I would also argue that any grammatical construction would have a grammatical function to perform, too, in addition to a possible function of adding an extra syllable in verse, especially isosyllabic verse in ME,

the poets had a considerable inventory of handy ‘fillers’: semantically dispensable monosyllabic words: *and*, *full*, *now*, *for*, *some*, the grammatically redundant ‘pleonastic’ *this*, *that*. Put it differently, although individual major class words retained their trochaic shape, in connected speech iambic cadences were sufficiently frequent and easy to construct; this permits an effortless ‘fit’ between the prosody of the language and the metre of verse.

In a similar way to Brinton (1996), I, therefore, argue that the function of *gan* and *can* to provide an extra syllable in a line of verse in ME could only be seen as promoting and spreading the use of this verb and its variant in verse – Brinton (1996: 71) observes that “[t]he requirements of verse alone cannot motivate the use of basic grammatical forms such as those of tense and aspect” (cf. Denison 1985: 45 on a similar argument on the function of *do* in verse). Furthermore, even though *gan* and *can* are said to have this function because they are meaningless, I argue that other items do not have to be meaningless in order to provide an extra syllable in verse to satisfy a particular type of metre. However, Inoue & Stokes (2009: 24) make a note

of the variant *con* occurring with the verb *mele* in infinitive in *Sir Gawain and the Green Knight* (l. 2295), which is an example of alliterative verse, as in (37) below. According to them, the verb *mele* is used in the preterite elsewhere in this poem. They argue, “[t]he only reason for using *con* here is to indicate a beat on *efte*, a semantically important word in this line. This is the second time that the Green Knight speaks and it vexes Gawain enormously that he does so. Gawain has asked him to strike without further unnecessary threats, and the Green Knight has said he will do so. Thus *con* is exploited to create the dissyllabic dip that confirms a beat on the narratively significant *efte*.” Nevertheless, the occurrence of this variant in *Sir Gawain and the Green Knight* verse would also suggest that *can* has a function, perhaps structural, in unrhymed, alliterative verse, whereas it has been shown that *gan* and *can* are infrequent in this kind of poetry. For example, Brinton (1996: 69) finds only one instance of *gan* in an alliterative poem.

- (37) *Pen* *muryly* *efte* *con* *he* *mele* *þe*
 Then merrily again did he address the
 mon *in* *þe* *grene*
 man in the green
 (*Sir Gawain and the Green Knight* l. 2295; Inoue & Stokes 2009: 24)

‘then he cheerfully addressed the man in the green again’

While the sources mentioned above concentrate on the function of *gan* and *can* to provide “an extra syllable or two to fill out a line of verse” (Brinton 1981: 120, cf. Brinton 1996: 83), a role of this verb and its variant to make the verb in infinitive available to provide rhyme has been considered by some sources, too; this role is in addition to the function of *gan* and *can* as a metrical filler or instead of it. For example, Taylor (1917: 574) argues, “[t]he periphrastic preterite, whether formed with *can*, *gan*, or *couth*, was merely a metrical device of the poet for obtaining the requisite number of syllables to fill the line, or for throwing the desired word into the rime position; usually the latter”. This kind of function to enable the verb in infinitive found with *gan* and *can* to provide rhyme is proposed by Koziol (1932: 132-133), Kerkhof (1966: 154, 158), Smyser (1967: 74), Terasawa (1974: 102), Tajima (1975: 434, 437-438) and Brinton (1981: 164). Statistical evidence also seems to compellingly back it: Tajima (1975: 433), for instance, notes about the verb in

infinitive found with *can*, “of the 80 instances found in the rhymed lines of *Pearl* and *Gawain*, 76, or 95%, are used for that purpose, while only 4, or 5%, occur without reference to rhyme”.¹¹² Next is Smyser (1967: 74), who maintains that in the *Tatlock-Kennedy* concordance, the verb in infinitive found with *gan* occurs at the end of the line in 73% of lines, whereas Brinton (1996: 70, 345) claims that this is the case for 94% of such cases in verse texts in the *HCET* and 66% in Chaucer’s *Troilus and Criseyde*. Overall, this kind of evidence is based on rhymed verse, with *gan* and *can* being infrequent in unrhymed, alliterative verse, as outlined in Section 6.2.3.

Smyser (1967: 74) also attempts to explain why the verb in infinitive found with *gan* occurs in rhyme position, claiming that the preterite, especially of weak verbs, is difficult to rhyme, resulting in the infinitive being used instead, “[t]he real advantage that *gan behold* had for Chaucer over *beheld* was a matter of meter and rhyme *Gan go, gan to go*, and (less common but not rare) *gan for to go* are convenient alternatives metrically to *wente*, and the infinitive of many verbs – particularly of weak verbs – are more easily rhymed than the preterites”. Smyser (1967: 74) supports his claims by referring to the distribution of the following weak verbs, whose forms of the infinitive and preterite are the same: *hente* ‘to lay hold of, seize, grasp’ (*OED* s.v. *hent* v. 1), *sette, stente (stinte)* ‘to extend, stretch out or set (a tent, sail, curtain, net, etc.) in its proper position’ (*OED* s.v. *stent* v.¹ 1) and *sterte* ‘to (cause to) make a sudden movement’ (*OED* s.v. *start* v. I). According to him, the expectation is that such verbs would rarely occur with *gan*; and the findings of his research indeed confirm that *hente* and *sette* feature twice in infinitive with *gan*, as opposed to in the preterite in rhyme position: 24 instances of *hente* and *sette*, 25 instances of *stente (stinte)* and 36 instances of *sterte*; this is in contrast to the distribution of the verb *cry*, which Smyser (1967: 74) finds occurring 21 times when used in infinitive with *gan* and six times in the preterite only (unlike the other verbs mentioned here, which are of OE origin, the verb *cry* derives from OF and enters the vocabulary of English at the beginning of the ME period; it also has the following preterite forms in ME: *cryde, criede, cryede* and *cried* *OED* s.v. *cry* v.). Therefore,

¹¹² Tajima (1975: 434) claims that “in the unrhymed alliterative poems of the Gawain-poet the *con-*periphrasis is very rare, occurring twice in the unrhymed alliterative lines of *Gawain*, 6 times in *Purity* and 3 times in *Patience*”.

Smyser's (1967) claims and evidence support the view that the role of *gan* would be to enable the verb in infinitive to occur in rhyme position; his claim that this is in particular because of weak verbs in infinitive is credible, too – I will come back to this point below.

While Smyser's (1967) proposals sound plausible in that they successfully attempt to explain, together with the supporting evidence, that the role of *gan* is to enable the verb in infinitive to occur in rhyme position in a line of verse, Fischer (1992: 266) suggests the role of the verb in infinitive is to enable this verb and its variant to occur in the preterite, “[a]nother interesting feature is that the construction occurs only in the past tense ... Smyser (1967: 4) explains this as follows”, providing the same explanation from this source, as the one in the preceding paragraph. She also observes that in (38) below, Chaucer resorts to symmetrical arrangements of the constituents, “it is likely that Chaucer used *gan* not just to put *bolde* in rhyming position, but also to put it in a symmetrical position to *ferre* with which it forms a contrast. Notice also the completely symmetrical ordering of all the other clause structures in these lines” (Fischer 1992: 266). However, I would argue that more evidence would need to be collected, other than the verb in infinitive occurring at the end of a line, to demonstrate symmetry of constructions with *gan*. In particular, therefore, Smyser's (1967) claims seem to work for former OE weak verbs to provide rhyme at the end of a line, such as *clappe*, as in (39) below, where the ME preterite forms would be *clapte*, *clappid*, *clapped* and *clapt* otherwise (*OED* s.v. *clap* v.¹). In the case of former OE strong verbs, on the other hand, Smyser's (1967) theory may provide an alternative rhyming possibility to the preterite form of such verbs (at least in most cases). Finally, Smyser's (1967) proposal might be useful in cases where there is the weakening of strong verbs, including in some dialects only, as in (40) below, this giving a useful possibility for the poet to provide rhyme at the end of the line. Overall, therefore, Smyser (1967) argues that the role of *gan* is to allow both weak and strong verbs to occur in rhyme position, with his data provided in the previous paragraph supporting this. However, if this were the only function of *gan*, there would be an expectation to see this verb rather frequently in verse, whereas the distribution of both is rather “skewed” in ME texts (Brinton 1996: 69).

- (38) *For with that oon encresede ay my fere,*
For with that one increased always my fear,
And with that other gan myn herte bolde,
And with that other began my heart embolden,
That oon me hette,
That one me hot,
That other dide me colde
That other did me cold
 (Chaucer *Parliament of Foules* 143; Brinton 1983: 237; translation in the original)

‘for with that one my fear increased, and with that other began my heart to be bold; that one made me hot, that other made me cold’

- (39) *The lyones fonde thay in hir dene,*
The lioness found they in her vale,
A knave childe laye sowkand hir then
A male infant lay sucking her then
And gan with the lyones to playe.
And began with the lioness to play.
Umwhile the childe sowkede hir pappe,
Sometimes the child sucked her breast,
Umwhile gan thay kysse and clappe;
Afterwards did they kiss and embrace;
 (1420-1500 *Octavian* l. 442-446; TEAMS)

‘They [the men] found the lioness in her vale, a male infant lay sucking her then and began to play with her. Sometimes the child sucked her breast, afterwards, they kissed and embraced;’

- (40) *Dis Kyng had with hym a man*
This king had with him a man
Pat he mykell avaunce gan
That he much advance did
And made hym of gret might
And made him of great might
 (a1425 *Four Daughters of God* st. 7; MED s.v. *avauncen* v. 3a; Visser 1969: 1574)

‘This king had a man with him who he promoted much, and made him of great might’

Evidence in support of the metrical and rhyming function of *gan* and *can* has also been offered on the basis of the similarity of the distribution of this verb and its variant to that of periphrastic *do*. For example, Mustanoja (1960: 602, 614) argues,

“there is good reason to assume that the use of the periphrastic *do* is rather similar to that of *gin* in the periphrastic preterite ... Like *gin*, the periphrastic *do* seems to be used in poetry largely to allow the poet to place the infinitive of the main verb at the end of the line”, whereas Smyser (1967: 81-82) claims, “[t]he meaningless *gan*-form in Chaucer serves metrical and rhyming convenience; paradoxically the *do*-form, though meaningful, may often do likewise, for the causation may be stated with a *do* (or *make* or *let*) or it may be left to inference.” Both Kerkhof (1966) and Fischer (1992) have made similar points. For example, Kerkhof (1966: 152, 154, 158, 159) maintains that the verb in infinitive found with *gan* occurs in rhyme-end position, just like the verb in infinitive found with periphrastic *do*, adding further on that “[*gan*] in this function [i.e. in prosodic function] competes with to do in poetry, which auxiliary in Chaucer is far less common than gan.” Further support for periphrastic *do* having a similar role comes from Ellegård (1953: 146, 208), who claims that it was “a peculiarity of the poetic diction, belonging to the paraphernalia of the verse-maker’s craft”, and Fischer (1992: 268), who shows that in “most of the early examples”, it occupied unstressed position in verse, thus suggesting that periphrastic *do* had a non-lexical function there.

However, while the above claim about the similarity in distribution between *gan* and *can* and periphrastic *do* may be supported in terms of the distribution (verse) and structure (infinitive) of these verbs, essentially, there appears to be a difference between these constructions in terms of meaning, i.e. meaningless *gan* and *can* and meaningful *do*, as argued by Smyser (1967). In this respect, *gan* and *can*, as well as *do* would not necessarily be in free variation, as suggested by the above sources in relation to the distribution and structure of these verbs. Furthermore, there might also a problem related to the timing of the occurrence of *gan* and *can* and periphrastic *do*, especially in the northern varieties of ME and EModE. For example, Nurmi (1999: 233-234) shows that there were very few instances of causative and periphrastic *do* in the north in the 15th century (one instance of causative *do* in prose and two instances of periphrastic *do* in verse). Nevertheless, the parallel drawn between *gan* and *can*, and *do* is that they all seem to have been made use of in verse for metrical purposes, with no necessity for them to have been used by the same poet for that

parallel to hold. Finally, it will be shown in Section 6.3.1.1 that, for example, aspectual and modal (auxiliary) verbs occurring in verse texts in the NME and ESc data sets can be seen in a metrical line with their verbs in infinitive at the end of it, although this does not necessarily mean that *gan* and *can* could not be used in a context where the use of this verb and its variant is not appropriate.

In addition to comparing the metrical function of *gan* and *can* to that of periphrastic *do*, an alternative proposal put forward in the literature concerns the similarity of *gan* and *can* in ME and Scots to historical present. Mustanoja (1960: 614), for example, suggests that this verb and its variant are interchangeable with historical present, arguing that there is a degree of competition between these constructions in ME, “[t]he intensive periphrastic preterite, used in vivid descriptions of past events, frequently occurs side by side with the historical present ... [i]n fact the two constructions seem to compete with one another” (cf. Kerkhof 1966: 154, Smyser 1967). Furthermore, the function of historical present being used for the sake of metre is also noted by Visser (1964: 135-142; 1972: 705-726). “[e]ven more common than the use of the present tense to past happening for the sake of rhyme is its use for the sake of metre” (Visser 1964: 138). However, unlike Mustanoja (1960: 614), Visser (1964) makes no comparison between the use of the verb *gan* and its variant *can*, and historical present. Therefore, I would disagree with suggestions that the function of *gan* and *can* is like that of historical present. Although related to aspect in that the function of historical present is “to impart vividness to the dead past”, to use Lloyd’s (1979: 89) expression describing historical present in general, i.e. to focus on complete actions or events, the historical present differs from *gan* and *can* in that it narrates complete actions or events – by means of any lexical verb – in the present, whereas the function of *gan* and *can* is to denote complete actions or events in the past. Therefore, the overlapping of the same function of historical present and *gan* and *can* is far from complete and proven, and so appears to be an issue here in that the use of historical present and the distribution of this verb and its variant are different, but complementary over time, genre (author) and/or dialect (Brinton 1996: 71).

6.3.1.1 *Literary Environment of Gan and Can, and Rhyming Function of Verb in Infinitive in Northern Middle English and Early Scots*

With proposals for the metrical function of *gan* and *can* in ME and Scots being difficult to prove, as outlined in the preceding section, I argue here that the evidence presented in Section 6.2 for this verb and its variant points to them being an integral part of verse, rather than being specifically used in verse to fill in a line with an extra syllable. The findings also suggest that, more specifically, the environment in which *gan* and *can* are found initially involves octosyllabic couplets in narrative texts in the NME and ESc data sets, as discussed in Section 6.2.2. However, *gan* and *can* are also used in other types of stanza in narrative texts in the NME data set in Period II, as such stanzas were also used in romances beside the octosyllabic line – in his corpus of around 50 medieval romances, Pearsall (1988: 16), for example, identifies 19 tales written in octosyllabic couplets, whereas 25 composed in tail-rhyme stanza, in addition to another six closely related to tail rhyme. Therefore, I argue that the use of *gan* and *can* in the NME and ESc data sets is linked to narrative texts written in the prescribed type of stanza (initially, octosyllabic couplets), with texts characterised by the octosyllabic line being defined as “more prosaic, realistic, historical and martial”, whereas those characterised by the tail-rhyme, “more emotive, more concerned with love, faith, constancy and the marvellous” (Pearsall 1988: 16).

While *gan* and *can* occur in narrative verse texts that are initially composed in the octosyllabic line in the NME and ESc data sets, this verb and its variant perform a function in verse in these data sets, too – rather to fill a line of verse, they make the verb in infinitive available to provide rhyme. The claim for this function is supported by the fact that nearly all or all such instances of the verb in infinitive are found in rhyme position, as outlined in Section 6.2.3, with some variation with respect to the data sets and periods. However, further research of some verse texts in the NME and ESc data sets reveals that in many cases, the verb in infinitive is also found at the end of a line when occurring with verbs functioning as causative and auxiliary verbs, as well as modal (auxiliary) verbs, alongside instances containing *gan* and *can*. This, in

particular, can be observed in *Octavian*, where not only is it possible to find a verb in infinitive in rhyme position with **gan**, but also with the modal (auxiliary) *sall*, as in (41) below. A similar point can be made for *Sir Eglamour of Artois*, where the modal (auxiliary) verbs *wyll* and *wolde*, as well as the causative *gart* are used in the same stanza with the verbs in infinitive occurring in rhyme position, as in (42) below. However, just because verbs in infinitive can be found in rhyme position with other verbs does not say anything about the issue of whether **gan** and **can** had a rhyming function or not (if infinitives in general were rarely found in rhyme position, that would indicate that they were difficult to rhyme, which is clearly not the case here). Therefore, to argue that **gan** and **can** would be doing nothing special in comparison to other verbs found with verbs in infinitive, it would need to be shown that verbs in infinitive would end up in rhyme position as frequently as is the case for verbs in infinitive with **gan** and **can**. Nevertheless, the occurrence of *beginnen* in the NME and ESc data sets suggests that this is not the case, as shown in Section 6.2.3.

- (41) *Than in his armes he gan hir folde*
 Then in his arms he did her embrace
And all his sorow he to hir tolde
 And all his sorrow he to her told
 ...
For fay we sall hythen fownde,
 For foe we shall hence go,
And I ne wote how this land sall fare
 And I not know how this land shall fare
Bot lyfe in werre and in kare
 But live in war and in care [sorrow]
 (1420-1500 *Octavian* l. 61-63, 66-68; *TEAMS*)

‘He then embraced her in his arms and told her all his sorrow ... we shall hence go, and I do not know how this country will prosper, but live in war and in sorrow.’

- (42) *The knyght was both bold and stronge*
 The knight was both bold and strong
Therfor the lady loved hym longe.
 Therefore the lady loved him dearly.
Lestenyth, I wyll you tell.
 Listen, I will you tell.
Syr Eglamour, he gart crye
 Sir Eglamour, he did call
 ...

What manere man her wold have,
 What kind man her would have,
So sore buffettys he hem gave
 So sore blows he them gave
For ever he gart them to dwelle
 For ever he made them to dwell [remain]
 (1420-1500 *Sir Eglamour of Artois* l. 37-40, 46-48; *TEAMS*)

‘The knight was both bold and strong and so the lady loved him dearly. Listen, I will tell you. Sir Eglamour, he intentionally made known ... What kind man would have her, he gave them so severe blows, for always he made them remain (i.e. killed them)’

6.3.2 Intensive-Descriptive or Stylistic Function

In addition to the metrical role of *gan* and *can* and/or the rhyming function of the verb in infinitive found with this verb and its variant described in the literature on the subject, it has been suggested that the function of *gan* and *can* is intensive-descriptive or stylistic in ME. Wuth (1915), Funke (1922: 8), Häusermann (1930: 18-22), Homann (1954: 389-398), Mustanoja (1960: 611-614), Kerkhof (1966: 156) and Terasawa (1974: 101-104) claim that in this role, this verb and its variant focus on the event or situation expressed by the verb in infinitive, thus isolating it from surrounding events or situations in the narrative. For example, Homann (1954: 389) argues, “Chaucer’s awareness ... leads him to utilize [*gan*] stylistically for distinctions in tempo, intensity and manner. And dramatically he lists the periphrastic preterit as a trenchant verbal subtlety for character behavior.” In a similar way, Mustanoja (1960: 611) maintains, “[*gan*] develops into a popular feature of ME narrative diction. Through frequent use it loses its ingressive colour more and more and occurs mainly in the intensive-descriptive function”. Quoting Homann (1954), Mustanoja (1960: 612) then continues that this function can be used whenever the poet faces a choice between this construction and the preterite, as in (43) below, where “*gan behold* seems to bring out in a vivid way the poet’s aesthetic experience”, and in (44) below, where “*beheld* records the mere fact”. He finds that “[t]he majority of instances consist of periphrastic preterites where the intensive-descriptive aspect makes itself more or less clearly felt or where, seemingly at any rate, the periphrastic preterite is used simply for metrical reasons”, as in (45) below (Mustanoja 1960: 613-614).

- (43) *This egle ... I gan beholde more and more,*
 This eagle ... I did behold more and more,
To se the beautee and the wonder
 To see the beauty and the wonder
 (Chaucer *House of Fame* II 21; Mustanoja 1960: 612)

‘I beheld ... this eagle more and more, to see the beauty and the wonder’

- (44) *Tho gan I loken under me,*
 Then did I look under me,
And behelde the eyrish bestes
 And beheld the airlike beings
 (Chaucer *House of Fame* II 456; Mustanoja 1960: 612)

‘Then I looked below me, and saw the demons of the air’

- (45) *How his auncestre Afrycan so deere*
 How his ancestor African so dear
Gan in his sleep that nyght to hym apere
 Did in his sleep that night to him appear
 (Chaucer *The Parliament of Fowls* 42; Mustanoja 1960: 614)

‘How his dear African ancestor appeared to him in his sleep that night’

Nevertheless, this intensive-descriptive function of *gan* in ME has attracted some criticism by being deemed as unconvincing, mainly because of subjective views where emphasis lies in a text, leading to over-interpretation. For example, Homann’s (1954) suggestions that the function of *gan* is intensive-descriptive or stylistic have been assessed by Visser (1969: 1572), who argues about it thus,

[t]his surely smacks of what in German is called ‘hineininterpretieren’; arbitrarily inventing motives for a writer’s use of a particular idiom of which motives the writer himself is not in the least aware. The above-mentioned grammarians in these suggestions overlook the fact that if their theories were correct the *gan/can* periphrasis would have had to be as common in prose as in poetry

Although Mustanoja (1960) advocates the intensive-descriptive function of *gan*, in a subsequent publication, he reviews his earlier stance,

[i]t seems apparent that the definition ‘intensive-descriptive use’ of *gan* (the term ‘descriptive’ comes from Funke) ought to be replaced by the more

neutral one ‘periphrastic’ or ‘metrical’ use. There is obviously reason to soften the tone of the statement beginning on p. 612, line 21, so that it runs ‘We have good reason to believe that in addition to metre and rhyme other stylistic considerations may have played a role in a good poet’s choice between the simple preterite and the *gan* periphrasis’. (Mustanoja 1983: 64)

Although not as harsh in their criticism as Visser (1969), other sources, too, point out that the intensive-descriptive or stylistic function of *gan* and *can* is, in general, difficult to prove. Tajima (1975: 436-437), for example, argues about this verb and its variant, “this kind of argument seems to be highly subjective, depending very much on the reader’s individual interpretation”. I also argue that the problem with the intensive-descriptive or stylistic function lies in the fact that an interpretation of the meaning of *gan* and *can* is purely derivable from the context, not taking into account the semantics of this verb and its variant (the effects of the Aktionsart of the verb in infinitive and durative or iterative adverbials, as well as other contextual information on the meaning of this verb and its variant). In other words, the intensive-descriptive or stylistic function is concerned with an interpretation of what the author of a text has probably intended, with a possibility of a subjective interpretation by the reader, bearing in mind that it is not possible to fully know authorial intent, especially when dealing with historical material. However, despite the criticism, Brinton (1988: 122) argues that this function of *gan* offers more than the metrical and prosodic ones, “I think that the function suggested by Wuth, Funke, Häusermann, and others, which they termed ‘descriptive’, has been misunderstood as ‘intensive’ or ‘emphatic’ ... I would like to suggest that ... it has a neutral narrative function”. Subsequently, Brinton (1996: 74) proposes textual and interpersonal functions of *gan* in ME verse, “when combined with recent work in discourse analysis, [the stylistic analysis of *gan*] provides an avenue for approaching the function of this periphrasis.” This function is discussed below.

6.3.3 Textual Function

Given the possibility of over-interpretation in the intensive-descriptive or stylistic function for *gan* and *can* in ME, a number of sources have advocated an alternative role for this verb and its variant in a verse text: for example, Brinton (1988) claims that *gan* mainly performs a textual function, in her later study (Brinton 1996) also

arguing for an interpersonal function of this verb, with the metrical function contributing to the spread of this construction in ME verse. Both studies by Brinton (1988; 1996) show an advantage of the textual function over the other ones, more specifically the difficult-to-prove intensive-descriptive function. More specifically, in both accounts, Brinton (1988; 1996) formulates her proposals on the basis of the framework of discourse analysis, incorporating recent work in this area, but also analysing this verb and its variant from the point of view of grammaticalization. In this respect, both examinations offer a more scientific, and so more objective, as well as integrated approach for identifying the role of *gan* in ME. Their analysis is based on an interpretation of linguistic elements affecting the aspectual reading of this verb, more specifically the impact of the Aktionsart of the verb in the infinitive, durative or iterative adverbials, and nominal elements on the aspectual reading of the verb *gan*, but also the role of this verb in a verse text in terms of a text's internal structure. Furthermore, her proposals also highlight the universality of the approach of discourse analysis in that discourse analysis can be used to explain similar such cases, i.e. where the development of discourse markers is argued to be a case of grammaticalization, even when they perform a pragmatic function.

I now review the proposals by Brinton (1988; 1996) in more detail. As already referred to throughout this chapter and in this study, Brinton (1988: 121) argues that, primarily, *gan* performs a textual function in ME verse texts. Having discussed the cases where the context (the Aktionsart of the verb in the infinitive, durative or iterative adverbials, and nominal elements) could prevent an ingressive interpretation of *gan*, Brinton (1988: 122; 1996: 77) refers to this verb as having a “‘delimiting’ or ‘demarcating’” function in such cases, “‘indicat[ing] breaks or new stages in the narrative sequence, or delimit[ing] new situations in the text’” (Brinton 1988: 122). She continues that as new narrative situations can include those that happen by chance, i.e. unexpectedly, accidentally, or fortuitously (although one should assume that the author of a text knows where the narrative is going), it is not surprising to see the following types of verbs in infinitive in such contexts in Chaucer: verbs denoting motion, sight, noise, or displays of emotion. Overall, therefore, Brinton's (1988) proposal that the textual function of *gan* is the primary one follows from the

hypothesis of the unidirectionality of grammaticalization, whereby the meaningful textual function precedes the meaningless metrical function (Brinton 1996: 82), but differs from Fischer's (1992: 267) claim that "if *gan* was used as a stopgap, it could be easily turned to stylistic use". Furthermore, Brinton's (1988) proposal also highlights the dual nature of *gan*: it continues in its function as a marker of ingressiveness in some contexts, as discussed in Chapter Five, while taking on textual and interpersonal, but also, eventually, metrical, functions in others.

In order to understand the textual, but also interpersonal, functions of *gan*, it is necessary to introduce the notion of foregrounding, which is associated with conveying the poet's choice of content and its effect on the reader. The term 'foregrounding' in its textual aspect has been defined by Prague School structuralist Mukařovský as "deautomatization" of linguistic aspects of the text so as to "place in the foreground the act of expression ... itself" (Garvin 1964: 19, as cited in Boase-Beier 2011: §7.2). This term is different from the notion of backgrounding, which refers to "[t]he part of a discourse which does not immediately and crucially contribute to the speaker's goal, but which merely assists, amplifies, or comments on it" (Hopper & Thompson 1980: 280, as cited in Brinton 1996: 45). Since *gan* has been associated with foregrounding (Brinton 1996: 49), as opposed to backgrounding, I provide below a number of definitions of the former only so that the reader will familiarise themselves with its characteristics (for exhaustive accounts on foregrounding, cf. also van Peer 1986, Douthwaite 2000, Jeffries & McIntyre 2010 and Leech 2013 amongst others). Douthwaite (2000: 178, as cited in Leech 2013: 4; emphasis in the original), for example, defines foregrounding as a technique of changing our perception of the world,

[i]mpeding normal processing by showing the world in an unusual, unexpected way or abnormal manner is termed **defamiliarization**. Thus defamiliarization may be achieved by subverting the rule governing perception and behaviour. The **linguistic technique** employed in subverting the world in this manner is termed **foregrounding**.

Similarly, Simpson (2004: 50) defines foregrounding as "deviation from a norm", but also as "more of the same",

a form of textual patterning which is motivated specifically for literary-aesthetic purposes ... foregrounding typically involves a stylistic distortion of some sort, either through an aspect of the text which deviates from a linguistic norm or, alternatively, where an aspect of the text is brought to the fore through repetition or parallelism ... [regardless of the type] the point of foregrounding as a stylistic strategy is that it should acquire salience in the act of drawing attention to itself. Furthermore, this salience is motivated purely by literary considerations and as such constitutes an important textual strategy for the development of images, themes and characters, and for stimulating both effect and affect in a text's interpretation.

Van Peer, Zyngier & Hakemulder (2007: 5) maintain that the theory of foregrounding "certainly allowed (if not promoted) a separation of textual features from reader effects. This separation of textual features from the effects of such features on readers may seem obvious today to some of us, but it is still not self-evident, either in stylistics or in literary studies in general." Quoting Hopper (1979a; 1979b) and Hopper & Thompson (1980), Brinton (1996: 45) provides a number of characteristics of the foreground as a feature of textual organisation: a) the foreground is usually seen as "the actual storyline of the narrative", forming "the main eventline of the narrative, as its 'skeleton' or 'backbone'", on the basis of which a summary of the story can be made; b) the events that are significant to the narrative are "asserted or represented", as opposed to being "reported or summarized"; and c) the events are given in chronological order, each event being presented as complete before the description of the next one is provided; this order is "iconic", i.e. representing the order of events in the world that is contained in the description. Overall, therefore, since texts are uneven in their organisation, as a feature of textual organisation, foregrounding is characterised by variability.

Either content or formal criteria can be used in order to describe foregrounded material. In relation to content criteria, Brinton (1996: 47), incorporating material mainly from Dry (1992) and Fleischman (1985; 1990), but also others, outlines some recognisable features of foregrounded clauses: a) the most salient characteristic of such clauses is "their temporal sequentiality", which, in turn, comprise the main story, or its backbone; b) foregrounded clauses are also "casually important" for the storyline, as they move the time of the narrative forward; c) such clauses concern stories around events or situations associated with humans and so are of human

interest and significance; and d) foregrounded clauses are of “thematic importance”, as they contribute to the overall comprehension of the story. While the content criteria deal with the relationship between foregrounded clauses and the importance of the material, formal criteria concern the relationship between foregrounded clauses and grammatical features. Brinton (1996: 49) provides a list of formal criteria characterising foreground clauses. A prototypical clause is, therefore, a main clause in the affirmative containing a human agent. Furthermore, this human agent focuses on a punctual event or situation (achievement, accomplishment), in which case it denotes the notion of perfectivity, i.e. completion of this event or situation. Finally, there is an overlap of grammatical and semantic features of grounding: “perfective events are interpreted as temporally ordered and casually related, while kinetic, agentive, voluntary events are of higher human interest” (Brinton 1996: 49).

Relying on proposals by Halliday (1970; 1979), Brinton (1996) provides linguistic and textual evidence to support her claim for both functions for *gan* in Chaucer’s *Troilus and Criseyde*; she maintains that this verb is: a) unevenly distributed, which relates to “the amount of narrated action as opposed to reported speech or thought in each book”, adding that the highest proportion of this verb occurs during the actions bringing Troilus and Criseyde together, while the lowest proportion of it is found in a chapter which discusses the imminent separation of these two characters; b) found in declarative foregrounded clauses; in such contexts, it is active or perfective, and the subject of it is often a main participant in the narrative, while the verb in infinitive denotes a dynamic, telic event responsible for carrying the narration forward; c) linked to “salient or important turns in the course of events”, occurring in multiples as opposed to regular intervals; and d) most commonly found with temporal as well as resultative and casual AdvPs; it also frequently occurs in clauses following a *whan*-clause; and it is found with manner AdvPs (Brinton 1996: 75, 76-77). Overall, therefore, the occurrence of the verb *gan* in some of the above-mentioned contexts is typical of foregrounding.

Brinton (1996: 77) also identifies the following contexts in which *gan* occurs in Chaucer’s *Troilus and Criseyde*: a) “changes in the time of day”; b) “changes in

scene or cast”; c) “introduction to or conclusion of speeches”; d) “character internal changes such as resolutions or responses”; e) “turns in the general course of events”; and f) “fortuitous occurrences”.¹¹³ Brinton (1996) argues that all the sentences containing instances of the verb *gan* in the above work amount to a synopsis of the plot, but these types of contexts may vary by texts used. All in all, the role of this verb in providing a synopsis, together with the linguistic and textual evidence outlined in the preceding paragraph, constitutes good evidence in support of the foregrounding function of *gan* in Chaucer’s *Troilus and Criseyde*. Brinton’s (1996: 80-81) findings lead her to conclude about *gan* in this work,

[t]he course followed by *gan* conforms to the directionality of semantic/pragmatic change observed in lexical items undergoing grammaticalization ... namely, from propositional to textual and interpersonal meanings. Whereas aspectual meanings belong within the propositional domain, turn-of-events meaning belongs to within the textual domain. Furthermore, it seems that an interpersonal meaning develops from the textual meaning. Since situations marked by *gan* represent new or important stages in the plot development and sometimes happen suddenly, unexpectedly, or fortuitously, *gan* may acquire the interpersonal (emphatic, intensive) meanings mentioned by Funke and his followers. These meanings represent an increase in subjectivity ... as the speaker marks in the text places which he or she considers important. In discourse analytic terms the emphatic/intensive function is better understood as ‘internal evaluation’; that is, *gan* is a ‘means by which the narrator makes the story interesting, highlighting the relative importance of the various narrative events’

Overall, Brinton’s (1996) proposals on the textual and interpersonal functions of the verb *gan* are appealing, as they take advantage of textual information (above and below sentence level) in an attempt to provide an objective view of the function(s) of *gan* in context. Additionally, Brinton’s (1996) analysis can be applied to other

¹¹³ Brinton (1996: 78-79) also offers comparative evidence of other constructions to support her claim about the textual function of *gan*: a) the existence of a similar type of function in periphrastic *do* and historical present (cf. Section 6.3.1); b) the existence of a similar function in other ingressive aspectualizers: i) OE *onginnan*, which “signal[s] the beginning of new and significant episodes”, alongside motion verbs and perception verbs (Richardson 1991a: 50); and ii) ME *com* and infinitive or present participle in *Kyng Alisaunder* and in *Libeaus Desconus*, this verb being replaced by *gan* and *can* in a later MS of the same work, but not due to corruption or oral mishearing (Terasawa 1974: 98-99); c) ME *comsen*, whose use, according to Brinton (1996), suggests similarity to *gan* in that it is “usually, if not always, simply stressing the fact that something is actually done or occurs; e.g. *comsede to leden*, did guide; *~for to fade*, did fade; *~grete*; did greet, greeted” (*MED* s.v. *cōmsen* v. 2, cf. Visser 1969: 1375-1376); d) late OE and ME *take* and *go*, which “more plausibly express textual rather than ingressive meaning” because of “aspectually incompatible elements” with which these verbs are found (Brinton 1996: 78); and, finally, e) ME *break* and *burst (out)* “appear to have had a textual function when occurring in hendiadic [sic] structures” (Brinton 1996: 78-79).

constructions, as it is compatible with discourse analysis in general. Finally, I would argue that Brinton's (1996: 80-81) proposal that *gan* performs a textual function is advantageous from the grammaticalization point of view: a) this proposal deals with a change in *gan* from propositional to textual and interpersonal meanings; b) it assumes that the propositional domain is associated with aspectual meanings, whereas the textual domain consists of "turn-of-events" meaning;¹¹⁴ and c) it follows from the fact that the interpersonal meaning derives from the textual meaning. In this respect, Halliday's (1970; 1979) proposals applied by Brinton (1996) in her study provide the basis of the textual and interpersonal functions, as these proposals see the textual function as establishing a degree of coherence in a discursive interaction, whereas the interpersonal function of language as exchange, creating relationships. In other words, the interpersonal function presents the language user with options to use in a particular speech role, whereas the textual function uses such options in a language to create texts (Brinton 1996).

6.3.3.1 *Textual and Interpersonal Functions of Gan and Can in Northern Middle English and Early Scots*

Following Brinton's (1988; 1996) approach to the understanding of the role of *gan* in ME texts, I show here that the verb *gan* and the variant *can* in the NME and ESc data sets also perform textual and interpersonal functions. Since Brinton (1996: 77) refers to this role of *gan* as a narrative device in foregrounded clauses in Chaucer's *Troilus and Criseyde* as "'delimiting' or 'demarcating'", I use the following categories in demonstrating the defining role of this verb and its variant in the narrative in the NME and ESc data sets: a) significant events or situations¹¹⁵; and b) changes in behaviour of some character(s) (despite there is a degree of overlap between these

¹¹⁴ While Brinton (1996) looks at the role of *gan* only in Chaucer's *Troilus and Criseyde*, Richardson (1991a) investigates a textual function of *gan* and *can*, but each in a different text. More interestingly, Richardson (1991a: 89) finds that there is also a difference between the uses of *gan* in *Troilus and Criseyde* and *can (con)* in *Sir Gawain and the Green Knight*, "[t]he main difference between *gan* in *Troilus and Criseyde* and *con* in *Sir Gawain in the Green Knight* is its slightly different structural function; the former delimits events while the latter, much like Old English non-perfective forms, delimits episodes".

¹¹⁵ Since both events and episodes are referred to in this study, their definitions are provided here. According to Brinton (1996: 41), 'episode' is located between the event and the story in the level of discourse structure and is sometimes referred to as 'scene' in medieval narratives. In a short text, an episode refers to a semantic unit or its surface manifestation, namely paragraph. In a longer text, however, an episode may consist of a number of paragraphs.

two categories). Although Brinton (1996: 77) identifies more categories, I argue that her categories “introduction to or conclusion of speeches”, “turns in the general course of events” and “fortuitous occurrence” could be subcategorised as ‘important events or situations’, although this may not always hold for “introduction to or conclusion of speeches”, unless they are particularly important to the narrative, this appearing to be the case here, given the frequency with which *gan* and *can* occur in such contexts in the narrative). In what follows, I, therefore, show that: a) the author of a text, through the narrative voice, draws the reader’s attention by the use of the *gan* and *can* to contexts of some great significance, usually religious, given the types of texts in which this verb and its variant are found; by doing this, he or she provides a synopsis of a story (Brinton 1996: 76-78); and b) by isolating such events or situations and/or behaviour of some character(s) from the narrative, the author of a text, through the narrative voice, exercises a didactic role by stressing the importance of them by employing *gan* and *can*, as and when necessary.

When performing the delimiting function (cf. Brinton 1996: 77), one would expect to see *gan* and *can* occurring frequently in a sample. But, the results of this study show that the frequency of occurrence of this verb and its variant varied in the NME and ESc data sets. Although it appears that there is no clear pattern, some texts are clearly higher in the frequency of occurrence of them than others, as shown in Section 2.4.1.1.1. This skewed distribution of *gan* and *can* appears to be related to what Briton (1996: 77) identifies as a textual function of *gan* in Chaucer’s *Troilus and Criseyde*, being “associated with salient or important turns in the course of events”, “generally occur[ing] in multiples, clustering around significant events” (Brinton 1996: 77). Therefore, both the high frequency of *gan* in texts like *Sir Eglamour of Artois* and *Octavian* (although not in relation to the other verse texts used in this study, but in relation to other sources) and the distribution of *gan* and *can* in verse only (including verse written as prose, but also in declarative clauses, as part of the narrative, and rarely in direct speech) suggest that this verb and its variant perform textual and interpersonal functions in the NME and ESc data sets, too. However, Richardson (1991a) shows that ME poets manipulated aspect for narrative purposes, with both ingressive aspect and the present perfect signalling episode boundaries in

Canterbury Tales and *Sir Gawain and the Green Knight*, as well as **gan** marking the main event lines, whereas **do** being responsible for highlighting peak events in *King Horn*. Therefore, in addition to **gan** and **can** in the NME and ESc data sets, I accept that other constructions, as suggested in the literature (cf. Richardson 1991a, Brinton 1996), could perform textual and interpersonal functions in ME, and so in the NME and ESc data sets. This could also potentially account for the varied frequency of occurrence of **gan** and **can** in the data sets.

The findings of this study showed that in terms of meaning (Chapter Five), **gan** and **can** in the NME and ESc data sets undergo semantic bleaching, although the results also revealed a number of instances of this verb and its variant characterised by opacity and persistence. Therefore, I argue that while the gloss of ‘began to’ is still appropriate in some contexts of this verb and its variant, the meaning of **gan** and **can** is weakened in many others. When this verb and its variant develop a textual function, they stop being associated with ingressive aspect. I maintain that this constitutes a shift from ingressive with reference to a situation or event on its own to a situation or event considered in relation to other actions in the context. This kind of shift is also characteristic of the process of grammaticalization, given that this process involves changes in the morphological properties of this verb and its variant, namely morphological fixation and phonological reduction (Chapter Three), as well as reanalysis and analogy (Chapter Four). In this textual role, therefore, **gan** and **can** are associated with events and/or behaviour of characters in the narrative in the NME and ESc data sets. In terms of interpersonal function, on the other hand, the author of a text, through the narrative voice, highlights important narrative events by using **gan** and **can**, with this function being subjective (addressee-focussed or orientated), as the author provides emphasis to events or situations and/or behaviour of characters he deems significant in the narrative (cf. Brinton 1996). Evidence in support of these functions is presented below. However, throughout the subsequent sections, I do not refer to the other function of **gan** and **can** in the NME and ESc data sets, namely to enable the verb in infinitive to provide rhyme, as this function of this verb and its variant has already been discussed in Section 6.3.1.1.

6.3.3.1.1 Significant Events or Situations

An example of *gan* performing textual and interpersonal functions can be found in *Octavian*. In this work, this verb appears to signal significant episodes (I use the terms ‘episode’ and ‘scene’ interchangeably here) in the narrative (textual function), occurring either at the start or at the end of a scene, thus marking points of transition between them (only the first few are presented here). For example, the lady notices that the Emperor is worried about not having an heir, as in (46) below; the Emperor tells the lady about his sorrow, as in (47) below; the Emperor is thankful to God for what He sent and so before the day begins, he asks the priest to offer mass in thanksgiving; he finds his mother there praying, too, and expresses gladness that the Empress did not die in childbirth, as in (48) below; the Emperor feels sorrow as a result of finding out that the two male children are not his, but fathered by the cook’s boy, given that the Emperor is unable to have children of his own, as in (49) below; as the Emperor kills the cook’s helper in revenge, the lady has a vision in her sleep about a dragon taking away her children, as in (50) below. By using the verb *gan*, the author of this text, therefore, provides a synopsis of the plot of the story (cf. Brinton 1996). In addition to the textual function signalling different parts in the narrative, which comprise a synopsis of the story, as outlined above, the narrator also appears to emphasise these events or situations, through the use of *gan*, as he sees them significant for the story (interpersonal function).

- (46) *Bot when the lady that gan aspye,*
 But when the lady that did discover,
All chaunged than hir bryghte blyee
 All changed then her bright face
And scho syghede full sore.
 And she sighed full sore.
Scho felle hir lorde one knees agayne,
 She fell her lord on knees again,
And of his sorow scho gan hym frayne,
 And of his sorrow she did him enquire,
And of hs mekyll care.
 And of his great sadness.
 (1420-1500 *Octavian* l. 49-54; TEAMS)

‘But when the lady discovered that, her bright face changed and she sighed deeply. She fell on her knees before her lord and enquired about his sorrow and great sadness.’

- (47) *Than in his armes he gan hir folde*
 Then in his arms he did her embrace
And all his sorow he to hir tolde
 And all his sorrow he to her told
And all his hertis wonde.
 And all his heart’s wound.
 (1420-1500 *Octavian* l. 61-63; *TEAMS*)

‘He then embraced her in his arms and shared with her all his sorrow and problems.’

- (48) *Erly are the daye gan sprynge,*
 Early before the day did begin,
He did a pryste his messe to synge;
 He did a priest his mass to sing;
His modir thore he fand.
 His mother there he found.
“Sone,” scho said, “I am full blythe
 “son,” she said, “I am very happy
That the Empryse sall haf hyre lyfe
 That the Empress shall have her life
 (1420-1500 *Octavian* l. 100-104; *TEAMS*)

‘Early before the day began, he made a priest say mass; he found his mother there. “Son,” she said, “I am very happy that the Empress did not die in childbirth’

- (49) *A sorowe there to his herte gan goo*
 A sorrow there to his heart did go
That wordis moghte he speke no moo
 That words might he speak no more
But yod awaye full still
 But went away full still
 (1420-1500 *Octavian* l. 118-120; *TEAMS*)

‘He felt such sorrow in his heart that he could speak no words, but went away quietly’

- (50) *Ay lay that lady faste and slepee [sic]*
 As lay that lady fast and sleep
A dolefull swevenynge gan scho mete;
 A doleful vision did she dream;
Scho was a wofull wyghte.
 She was a woeful creature.

(1420-1500 *Octavian* l. 160-163; *TEAMS*)

‘As the lady lay sleeping, she had a disturbing vision; she was a woeful creature.’

As *gan* may be used to isolate and highlight significant stages in the narrative, occurring either at the start or the end of events or actions, it is also possible for this verb to refer to one event or situation within a single scene. For example, in *De fine Ade & oleo misericordie* in *Cursor Mundi*, the narrator presents two contrasting characters: Adam, who is *nine hundret yere* ‘nine hundred years old’ and *vn-ferre* ‘infirm’, and Seth, *his sun* ‘his son’. Adam asks his son to go to the gate of paradise, his abode before the Fall in the biblical account of the Creation, to tell the angel that he is *vnferre*, *For [he has] liued so mani a yere, Ai in strijf and soruuing stad, þat o [his] lijf [he is] al sad* ‘infirm, for he has lived so many years, always in strife and sorrowing place, that of his life he is all sad’. Next the reader is presented with a narrative description containing *gan*, whereby the path found by Seth marks him the way towards the gate of paradise, as in (51) below. In this instance, the narrator changes the action in the line of the narrative (textual function), referring to the path showing to Seth the way to the gate of paradise. In Christian terms, the act of showing the way to paradise may be metaphorically interpreted as denoting one’s journey to salvation, this being emphasised by the author in his didactic task by means of *gan* (interpersonal function).

(51) *Þe slogth he fand þat him gan wiss*
The path he found that him did mark
Tilward þe zate of paradis;
Toward the gate of paradise;
(1350-1420 *Cursor Mundi* p. 82; *HCET*)

‘He found the path that marked the way for him towards the gate of paradise;’

An observation that the author chooses to stress some significant events or situations by the use of *can* can be made in relation to *De ysaac filio abrahe* in *Cursor Mundi*, as in (52) below. Here, Laban promises his younger daughter, Rachel, to Jacob, but Jacob has not met Rachel before. This verb signals a turning juncture in the narrative,

emphasising the fact that he is to be her husband in return for seven years' service. Although seven years seems to be an excessive period of time to wait to marry someone, the importance here is on whether the characters can give themselves unselfishly for each other's happiness, i.e. whether they can grow to love each other. By focussing on that, the author is seen showing that infatuation can be self-centred. Jacob's decision to hastily marry Rachel has an important effect on the course of events or actions. This is because Laban puts Jacob to a test by tricking him into marrying Laban's older daughter, Lia or Lya, instead. As a result, Jacob serves another seven years for Laban in order to marry Rachel, his first choice. The reason why the author uses this verb here is to signal Rachel's realisation (textual function) that she is to marry Jacob without learning to intimately know him first, to which the author draws the reader's attention (interpersonal function). In this respect, *can* performs textual and interpersonal functions here.

(52) “*I am iacob, þi cosin nere,*
 “I am Jacob, your cousin near,
 For þi luue am i comen here.”
 For your love am I come here.”
 þai mai quen sco can vnderstand
 the womanwhen she did understand
 þat iacob suld be hir husband
 that Jacob should be her husband
 To laban tald sco new tiþand,
 To Laban told she new tiding,
 (1350-1420 *Cursor Mundi* p. 227; *HCET*)

“‘I am Jacob, your cousin, for your love I have come here.’ When the woman understood that Jacob should be her husband, she told Laban the new news,’

Alternatively, *gan* can be used to refer to a significant event or situation within the same episode. For example, in *Meditatio de passione Cristi in Richard Rolle of Hampole, Yorkshire Writers*, this verb is used to isolate two separate events in the narrative, i.e. the same episode, namely, Christ being beaten, as in (53) below, and Christ's death on the Cross, as in (54) below. However, it is the former use of *gan* that exercises textual and interpersonal functions. This verb occurs at the start of a narrative description of Christ's physical suffering and denotes the act of beating,

being spat on, receiving the thorn crown and being nailed to the cross; the latter, on the other hand, signifies the beginning of his death, i.e. blood pouring out of his wounds being an indication of impending death. By using *gan* in a non-ingressive sense at the start of the episode, the author attempts to emphasise the fact that the suffering and death of Jesus are the most far-reaching in their implications than any other event, be it Christ's birth or his resurrection. Overall, therefore, the author of the text epitomises Christ's role in man's salvation by drawing the reader's attention through the use of *gan* to the significant event of Christ's suffering presented at the start of the episode. This way, the author of the text fulfils a didactic role to teach the reader about the important events in Christ's history and their significance to man.

(53) *Ful fast þai gan hym dyng,*
 Full fast they did/began to him strike,
and at þe pyler swyng,
 and at the pillar swing,
& his fayre face defowlyng /
 & his fair face defouling /
with spittyng. /
 with spitting. /
 (1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 57;
ICoMEP)

'They quickly did/began to strike him and swung him at the pillar, and defiled his beautiful face with spitting.'

(54) *In five steads of his flesh þe blode gan*
 In five steads of his flesh the blood began
downe glyde, als stremes of þe strande
 down glide, as streams of the river
 (1350-1420 *Richard Rolle of Hampole, Yorkshire Writers* p. 57;
ICoMEP)

'In five parts of his body, the blood began to glide down, like streams of the river'

6.3.3.1.2 *Behaviour of Character(s)*

Both *gan* and *can* can also serve to indicate changes in the behaviour of characters, regardless of whether this involves the leading character (protagonist), or one of the main characters in the narrative. For example, in a part of *The Northern Homily Cycle*, the reader is presented with a description of a tired and thirsty Jesus, who is a

Jew, arriving at Jacob’s well and asking a woman, who is a Samaritan, for water. A conversation held between both of them reflects “the strained relations between Samaritans and Jews in the first centuries of the Common Era” (Knoppers 2013: 1). Furthermore, for Jesus to have a conversation with the Samaritan suggests that this episode is of some significance, as there are a number of Christian and universal truths being mentioned in the story, namely, that everyone can receive eternal life, that Jesus is the Messiah, and, most importantly, that all people are valuable to God, regardless of their race and sex. Overall, the function of *gan* in this part of the text, as in (55) below, is to mark the act of the Samaritan woman repenting all her deeds. By doing so, the author of this story appears to show to the reader that God loves everyone, including the Samaritan woman, i.e. an outcast; furthermore, salvation is only achieved by those who renounce or repent their deeds. This author, therefore, serves to stress the importance of the above by isolating the event from the context and emphasising its significance for the reader.

(55) *Pe woman þan hir dedes gan rew*
 The woman then her deeds did repent
And þus scho said vnto Jesu,
 And thus she said unto Jesus,
 (1350-1420 *The Northern Homily Cycle* P II p. 82; *HCET*)

‘The woman then repented her sins and said to Jesus in this way,’

The Parable of a King and his Four Daughters in *Cursor Mundi* is an allegorical description involving the king’s four daughters, namely *merci* ‘Mercy’, *sothfastnes* ‘Truth’, *rightwisnes* ‘Justice’ and *pees* ‘Peace’, a son and a faithless servant. Because of the servant’s misdemeanour, he is placed in prison and the four daughters beg the king to have him released. In this context, the use of *can*, as in (56) below, has a particular narrative juncture – it isolates the fact that *merci* is the first to ask her father to have the servant released and by using *can*, the author draws the reader’s attention to this part of the story. The role of *merci* having pity on the servant’s predicament is crucial here, as it can be linked to the fact that God himself is presented as merciful and gracious. References to God being compassionate are found in *Vices and Virtues*, on which *The Parable of a King and his Four Daughters* in *Cursor Mundi* is based. In this text, the balance between the four daughters is not

evenly held, with Truth maintaining that God is on Mercy’s side (Matthews 1962: 71-72). Therefore, the behaviour of *merci* as a character is meant to be seen as prompting a change in the behaviour of the king – the works of mercy are charitable actions. By doing so, the author appears to show that mercy, presented allegorically here, is an important virtue as God himself is associated with it. The narrative function of *can* is to emphasise the importance of *merci* feeling pity for the servant and, subsequently, asking the king, her father, to show him benevolence, too.

(56) *Quen merci sagh him suagat be*
 When Mercy saw him thus be
Of him sco can haf pite,
 Of him she did/began to have pity,
Sco moght hir forbere nathing
 She might her[self] restrain nothing [not at all]
 (1350-1420 *Cursor Mundi* p. 550; *HCET*)

‘When Mercy saw him be in this way, she did/began to have pity, she could not restrain herself in any way’

6.4 Summary

An overview presented in this chapter of the literature on the subject showed that *gan* and *can* mainly occur in verse in ME and Scots, although some sources argue that this verb and its variant can occur in prose, too. However, my findings demonstrate that *gan* and *can* are exclusively used in narrative verse in the NME and ESc data sets, including two instances of *gan* in what has been classed by the compilers of the *ICoMEP* as a prose text. This type of verse is characterised by alliteration, but applied for expressive or ornamental, rather than structural, purposes. The verb *beginnen*, on the other hand, is found in verse and prose alike in these data sets. Furthermore, my results also show that *gan* and *can* tend to be almost exclusively associated with a specific type of verse, namely the octosyllabic line, in the NME and ESc data sets in Period I, although instances of this verb and its variant are found in other types of verse in the NME data set in Period II, although this is not to say that this verb and its variant are only initially used in verse written in the octosyllabic line. Although the octosyllabic line was not the only type of line practised at the time, it was fashionable (K. Campbell 1975: lxxvi), as it was a

general-purpose type of cadence linked to ‘serious literature’, i.e. dealing with religious, didactic and historical matters (Damian-Grint 1999: 186) (the octosyllabic couplet was being replaced by the tail-rhyme stanza in romances from the mid of the 14th century; Pearsall 2011: 9). Finally, my results also reveal that the verb in infinitive found with *gan* and *can* in the NME and ESc data sets occur at the end of a line in most cases.

Given the fact that *gan* and *can* are a feature of verse in ME and Scots, and that the verb in infinitive is found in rhyme position, a number of sources have proposed various functions for *gan* and *can*, but such proposals have been met with varying degrees of acceptance and validity: metrical and rhyming, intensive-descriptive and textual. My analysis shows that in addition to a function of this verb and its variant to enable the verb in infinitive to provide rhyme (rhyming function) in narrative verse texts in the NME and ESc data sets, *gan* and *can* perform textual and interpersonal functions there: a) in the textual function, this verb and its variant have a demarcating role in the narrative, isolating events or situations, and changes in behaviour of some characters, from the rest of the narrative; and b) in the interpersonal function, on the other hand, this verb and its variant emphasise the importance of events or situations and changes of some characters; in doing so, the narrator performs a didactic role, conveying information about universal truths. By the content or formal criteria outlined in Section 6.3.3, the clauses with *gan* and *can* in the NME and ESc data sets are foregrounded, offering a better understanding of the textual and interpersonal functions of this verb and its variant in narrative verse. Overall, therefore, in line with principles of semantic and pragmatic change in the process of grammaticalization, *gan* and *can* perform textual and interpersonal functions in narrative verse in the NME and ESc data sets, with this verb and its variant also enabling the verb in infinitive to provide rhyme there.

Chapter Seven – Conclusions

7.1 Meeting Aims and Objectives

Instances of the verb *gan* and its variant *can* behave differently from the etymologically related verb *beginnen* in ME and Scots in that they mainly, or exclusively, occur with the plain infinitive and with a non-ingressive meaning. What is more, *gan* and *can* are found in narrative verse where they have a metrical, intensive-descriptive or textual function. All of this suggests that *gan* and *can* show further rightward movement towards auxiliation in ME and Scots than the verb *beginnen*. Existing studies mainly concentrate on the meaning and/or function of *gan* and *can* in verse (Wuth 1915, Taylor 1917, Beschorner 1920, Funke 1922, Mustanoja 1960, Kerkhof 1966, Visser 1969 and Brinton 1981; 1983; 1988, Mustanoja 1983 amongst others), whereas investigations by Brinton (1981; 1988; 1996), Ogura (1997; 1998; 2013) and Sims (2008; 2014) deal with the divergence in the development of this verb and its variant in terms of grammaticalization in ME. Studies by Los (2000; 2005), in contrast, address the grammaticalization of *onginnan* and *beginnan* with the plain infinitive in Ælfric’s works. Since no studies have been conducted on whether *gan* and *can*, as well as *beginnen* develop differently in terms of grammaticalization in NME and ESc, I addressed in this thesis whether: a) *gan* and *can*, as well as *beginnen* differ with respect to their morphological paradigms in the NME and ESc data sets, in view of what we know about grammaticalization and the development of invariant forms? b) these verbs differ with respect to their complements, in view of claims in the literature that the more grammaticalized variant takes the plain infinitive; and c) *gan* and *can* are a development from OE *onginnan* and OE and eME *aginnan*, originally expressing ingression, but shown in the literature to have undergone semantic bleaching in OE and in eME period?

With my research being influenced by recent proposals in the field of grammaticalization (Pagliuca 1994, Traugott 1995, Heine 1997, Hopper 1998, Croft 2000, van der Auwera 2002, Wegener 2002, Bybee 2003, Hopper & Traugott [1993] 2003, Brinton & Traugott 2005, Bybee 2006, Diewald 2010, Traugott 2010 amongst

others), I argued that the verb *gan* and its variant *can*, as well as the verb *beginnen* differ in the NME and ESc data sets, with the former two showing more divergence in their development towards auxiliation. From a historical perspective, then, *gan* and *can* undergo a number of mechanisms and processes typical of the grammaticalization process: decategorialization and associated morphological fixation, phonological reduction, analogy, reanalysis and semantic bleaching; *beginnen*, on the other hand, is subject to reanalysis, analogy and some degree of semantic bleaching, especially in the NME data set. I also argued for the grammaticalization of *gan* and *can* from the pragmatic point of view, where I showed that this verb and its variant have a number of functions in narrative verse in the NME and ESc data sets: to enable the verb in infinitive to provide rhyme (rhyming function) and to isolate events or actions, and changes in behaviour of some characters from the rest of the narrative and to emphasise their importance (textual and interpersonal functions). By using both perspectives on the grammaticalization of *gan* and *can* in the NME and ESc data sets, I showed aspects of a journey of this verb and its variant towards auxiliation.

Gan and *can* are different from *beginnen* with respect to their morphological paradigms by undergoing decategorialization and phonological reduction, although there is also evidence of analogical change. Regarding decategorialization, *gan* and *can* are morphologically fixed in the third person singular in the preterite, which is formed by root vowel changes; *beginnen*, on the other hand, is found in a range of finite and non-finite forms. Since *gan* and *can* only occur in the finite form, they display characteristics typical of auxiliaryhood, given that the lack of non-finite forms is a distinctive feature of modal (auxiliary) verbs, this property already characterising such verbs in IME (Warner 1993: 144-145). Consider all of these characteristics, I argued that there is a category change for *gan* and *can* in the NME and ESc data sets. With respect to phonological reduction, I maintained that it is possible to observe evidence of this process in a change from OE *onginnan* and OE and eME *aginnan* to *gan* in the NME and ESc data sets. However, I also claimed that there were two phonological processes occurring there, both of which as part of the process of grammaticalization: the first concerned the weakening of prefixes like

on, resulting in a change from OE *onginnan* to OE and eME *aginnan*, whereas the second one affected *gan*, especially when other mechanisms and processes are taken into account. Although not displayed in the NME and ESc data sets, there is also evidence of analogical change affecting *can*, whereby the variant *couth* emerges “[c]hiefly” in Scots (*OED* s.v. *can* v.²), but also in NME and SME (Visser 1969: 1579-1580). Finally, analogical change can be observed in relation to *beginnen* in the ESc data set only, where the variant form *begouth* is also used.

Gan and *can*, as well as *beginnen* show divergence with respect to their complements in the NME and ESc data sets. In consideration of claims in the literature that the more grammaticalized variant takes the plain infinitive, the former two shift towards auxiliatation; the latter, on the other hand, occurs with a range of infinitives, such as the (*for*) *to*-infinitive, the *at*-infinitive and the plain infinitive, although to a varying degree in the data sets and periods, with the (*for*) *to*-infinitive being the most frequent type here. While a number of verbs continued to take the *to*-infinitive as their complement in ME and Scots, the plain infinitive became restricted to a small group of verbs, amongst which were auxiliary verbs, including modal (auxiliary) verbs. I argued that in the NME and ESc data sets, the construction of *gan* and *can* with the plain infinitive arises out of a series of reanalyses, with evidence of this type of mechanism of change already operating on OE *onginnan* in Ælfric’s works, when this verb occurs with the plain infinitive (Los 2000; 2005). I also maintained that in the NME and ESc data sets, the construction of *beginnen* with infinitive develops out of reanalysis, but unlike in the case of *gan* and *can*, this verb also continues to function as a lexical verb. Therefore, in the case of *gan* and *can*, as well as *beginnen*, what was initially one category of verbs is now reanalysed as two, with the series of reanalyses in relation to the former being part of the process of grammaticalization. Furthermore, *gan* and *can* have some selectional restrictions on the type of infinitive that they take in the NME and ESc data sets; while restrictions on the type of infinitive with which *beginnen* co-occurs also apply here, gradually the (*for*) *to*-infinitive is becoming the pattern, as is the case for this verb in the NME data set, or is already the norm, as evidenced in the ESc data set.

An example of diachronic change, *gan* and *can* are also a development from OE *onginnan* and OE and eME *aginnan*, originally expressing ingression, but presented in the literature to have gone through semantic bleaching in OE and in eME period. In the NME and ESc data sets, the semantically bleached, or non-ingressive, meaning of *gan* and *can* occurs after the decategorialization and reanalysis of this verb and its variant (cf. Hopper & Traugott [1993] 2003), when a verb in infinitive is a punctual (achievement) verb. This type of meaning is different from the ingressive meaning associated with the Germanic *ginnan*, following the transition of the original sense ‘open up’ to ‘begin’ (*OED* s.vv. *begin* v.¹; *ongin* v.). Similarly, the occurrence of *gan* and *can* with other elements within a clause which are otherwise incompatible with ingressive aspect, for example adverbials of duration or iteration, also shows that the meaning of this verb and its variant is non-ingressive in the data sets under investigation here. Other contextual information also seems to help to establish the meaning of *gan* and *can*, especially in situations when the Aktionsart of the verb in infinitive can be difficult to classify. Furthermore, I also demonstrated in this thesis that a small number of instances of *gan* could still be used with the ingressive meaning, this behaviour of this verb being typical of a situation known as persistence. In other contexts, it was also possible to interpret *gan* and *can* as either non-ingressive or ingressive, especially when the necessary contextual clues were unavailable. The fact that a number of instances of *gan* and *can* in the NME and ESC data sets are characterised by persistence and opacity highlights the constraint between the source and the outcome of the change, namely the auxiliation of this verb and its variant.

My study also showed that *gan* and *can* perform a pragmatic function in narrative verse in the NME and ESc data sets (two instances of *gan* occur in verse written as prose, as part of the *Richard Rolle of Hampole, Yorkshire Writers* prose text), unlike *beginnen*, which occurs in both prose and verse. The type of verse in which this verb and its variant are found is mainly the octosyllabic line, but also other types of verse in the NME and ESc data sets, as they become more popular and widespread during the ME period. By occurring in verse, *gan* and *can* make the verb in infinitive available to provide rhyme, but also have textual and interpersonal functions, unlike

the verb *beginnen*, whose sole function is that of a marker of ingressive aspect. In the textual function, *gan* and *can* mark events or actions in the narrative and isolate changes in behaviour of some characters. In the interpersonal function, on the other hand, the author, through the use of this verb and its variant, emphasises the importance of the above, performing a didactic role. The former is a neutral function, whereas the latter is subjective. Finally, from the grammaticalization point of view, more specifically in line with the unidirectionality of grammaticalization, the textual and interpersonal functions of *gan* and *can* develop before the role of this verb and its variant to enable the verb in infinitive available to provide rhyme, which is viewed as meaningless (cf. Brinton 1996: 81-82).

My investigation, therefore, indicates that *gan* and *can* show a greater degree of grammaticalization in the NME and ESc data sets than *beginnen*. The grammaticalization process of this verb and its variant involves the following: a) changes in the underlying structure of the category of this verb and its variant with a punctual (achievement) verb in the plain infinitive (reanalysis and analogy), which are manifested by alterations of the surface structure, namely morphological fixation and phonological reduction; b) changes in the semantics of *gan* and *can*, characterised by a shift from ingressive to non-ingressive, this change being the outcome of semantic bleaching)¹¹⁶; and c) changes in functional domain in that this verb and its variant occur in narrative verse to make the verb in infinitive available to provide rhyme, with the other functions being textual and interpersonal. The process of grammaticalization of *gan* and *can*, therefore, denotes the rightward movement from grammatical (OE *onginnan* and OE and eME *aginnan* with the plain infinitive) to more grammatical along the cline of grammaticality. In this respect, non-ingressive *gan* and *can* with the plain infinitive show different characteristics in the NME and ESc data sets from ingressive *beginnen* with infinitive. However, given that a number of instances of *gan* and *can* can still be interpreted as either ingressive

¹¹⁶ The difference between non-ingressive and ingressive meanings which can be found amongst instances of *gan* and *can* is based on how the event or action is viewed: in relation to the ingressive meaning, as this event or action is presented as initiated, with the outcome of this event or action being interpreted as either complete or not complete; in relation to the non-ingressive meaning, on the other hand, the event or action is interpreted as complete, i.e. in its entirety, without references to temporal structure of this even or action.

or non-ingressive, it appears that the grammaticalization process of this verb and its variant is advanced, but, perhaps, not fully completed yet.

7.2 Regional Differences in the Distribution of *Gan*, *Can* and *Beginnen* in Northern Middle English and Early Scots

Despite the received view that the ‘English’ written in the six northern counties of England during the late 14th and the 15th centuries and the ‘English’ written in Scotland at the time refer to a common speech area (Murray 1873: 29), the 15th century was a period of linguistic divergence in these varieties and in those respective areas (Williamson 2002: 253, Kopaczyk 2013: 252). Therefore, this study shows that in the NME and ESc data sets, which are, to an extent, representative of the English of the north of England and the English of Scotland in the 14th and the 15th centuries, respectively, variation between both data sets concerns *gan* and *can*, as well as *beginnen*. For example, these verbs occur in the NME data set, with only *can* and *beginnen* being found in the ESc one. But, this varied distribution does not mean that *can* only occurs in ESc in general. Although Macafee (1992/93: §8.13) states that most poets used this variant in Scots, both Barbour and Douglas are an exception: in relation to Barbour, the occurrence of *can* appears to be MS-specific, with the Cambridge MS of this work using this variant, whereas the Edinburgh MS of *The Bruce* employs *gan* instead (Taylor 1917: 574). An instance of variation between the NME and ESc data sets also concerns the distribution of the verb *beginnen*, whose variant form *begouth* is recorded in the ESc data set only, in line with statements in the literature that this variant form of *beginnen* is only found in Scots.

Furthermore, this study shows that the occurrence of *gan* and *can* is uneven in the NME and ESc data sets. However, the skewed distribution of this verb and its variant is directly related to the type of text in which *gan* and *can* perform a number of functions, rather than either of the data sets used here. For example, narrative verse texts like *Octavian* and *Sir Eglamour of Artois* in the NME data set have much higher frequencies of occurrence of *gan* and *can* than any other verse text used in this study. What is more, while there is a link between the frequencies of occurrence

of this verb and its variant and narrative verse texts in the NME and ESc data sets, it cannot be said about *beginnen* that the distribution of this verb points towards this verb being specifically a feature of either prose or verse, with the distribution of this verb relying on the common units, such as sentence and paragraph. However, this study also demonstrated that there is a tendency for this verb to occur in prose, more so than in verse, especially in the ESc data set, but this could be down to accidental gap. Overall, therefore, in terms of dialectal differences, although *gan* and *can* appear to occur more frequently in the NME data set, this increased frequency is due to the type of verse texts used in this data set.

Finally, there is very little dialectal variation amongst instances of *gan* and *can* in terms of their complementation – the former is almost exclusively found with a verb in the plain infinitive in the NME data set, with only one instance of this verb found with the *to*-infinitive; the latter, on the other hand, exclusively occurs with the plain infinitive in the NME and ESc data sets. However, dialectal differences can be observed in relation to the complementation of *beginnen* in the NME and ESc data sets. For example, in the former data set, this verb occurs with a variety of infinitival complements, such as the plain infinitive, the (*for*) *to*-infinitive and the *at*-infinitive, in addition to functioning as an independent verb, whereas in the latter data set, it is found with the (*for*) *to*-infinitive only. Overall, therefore, there are dialectal similarities in terms of the complementation of *gan* and *can* in the NME and ESc data sets, showing a greater degree of grammaticalization of this verb and its variant in both data sets, in line with claims in the literature that the more grammaticalized variant takes the plain infinitive. However, there are some differences, albeit minor, in relation to the type of complementation found with *beginnen* in the NME and ESc data sets, showing that this verb continued to take the *to*-infinitive in both data sets.

7.3 Limitations of the Study

The conclusions provided in this chapter have been reached on the basis of the analysis of *gan* and *can*, as well as *beginnen* in the NME and ESc data sets only. An obvious limitation of this study, therefore, concerns the data sets used here. More specifically, there is a problem of the availability or lack of written material before

NME and ESc, respectively, in order to accurately investigate the extent of grammaticalization of *gan* and *can*, as well as *beginnen* in the NME and ESc data sets. The lack of this material means that we cannot be sure when exactly these verbs, where applicable, begin to grammaticalize in the northern varieties of English and in Scots, especially that the varieties in the north of England and in Scotland are traditionally seen as grammatically innovative when compared with those found in the south of England. Furthermore, because of the problem with the availability or lack of textual evidence, we are unable to establish whether these verbs grammaticalize at the same rate and to the same extent in the varieties used in the north of England and in Scotland, given that Los (2000; 2005: 88-99) claims that OE *onginnan* and OE *beginnan* already display auxiliation in Ælfric's works.

In addition, there is an issue with the amount and the quality of material written in NME and ESc, and so available in the NME and ESc data sets. This problem is often overlooked in the literature. A good example to illustrate this can be taken from Terasawa (1974), who lists a number of northern texts used for his analysis (*Awntyrs, Degrevant, Eglamour* and *Tottenham*). However, such texts are effectively only originally composed in a northern dialect, because the editions used by this source are de facto from subsequent MSS, characterised by a scribe's own dialect, which is not necessarily northern. That this would constitute a problem when carrying out any dialectal study is best illustrated with an example from the *MED* (s.v. *ginnen* v. 3b), where different MSS of *Cursor Mundi* contain different verbs in the place of *gan*, i.e. *cun*, *con* and *dud*. Another limitation of this study concerned the availability of a range of textual material, resulting in an imbalance, in the ESc data set especially, between texts that are classed as prose and those that are categorised as verse. Despite limitations, patterns of variation and change of *gan* and *can*, as well as *beginnen* in the NME and ESc data sets do indicate that the former two shift rightwards towards auxiliation, whereas the latter one remains no longer quite like a lexical verb, but not quite yet like *gan* and *can* in auxiliary uses. However, it is not possible to say more than this until further research into both NME and ESc is undertaken, as even larger, and more varied and diplomatically edited, corpora containing texts written in the above dialects become available.

Appendix One

Table 58

NME Texts from the HCET

	Text	Type of text	Word count
Period I	<i>The Northern Homily Cycle</i>	Homily (verse)	7,576
	<i>The Benedictine Rule</i>	Rule (prose)	2,555
	<i>The Pricke of Conscience</i>	Religious Treatise (verse)	6,018
	<i>Cursor Mundi</i>	History (verse)	10,521
	<i>Petition of Willyam Midylton</i>	Document (prose)	406
	<i>The Liber De Diversis Medicinis</i>	Handbook, Medicine (prose)	5,763
	<i>Dan Jon Gaytryge's Sermon</i>	Sermon (prose)	5,373
Period II	Rolle's <i>The Bee and the Stork</i>	Religious Treatise (prose)	546
	Rolle's <i>Prose Treatises</i>	Religious Treatise (prose)	6,324
	Rolle's <i>The Psalter</i>	Religious Treatise (prose)	13,001
	<i>The Wakefield Pageants</i>	Drama, Mystery Play (verse)	4,120
	<i>The York Plays</i>	Drama, Mystery Play (verse)	5,562

Table 59

NME Texts from the ICoMEP

	Text	Type of text	Word count¹¹⁷
Period I	Hilton's <i>Angels' Song</i>	Religious mysticism/Homily – Instruction, religious (prose)	2,413 (2,413)
	Richard Rolle of <i>Hampole, Yorkshire Writers (rollhor1.rtf)</i> , including <i>Meditatio de Passione Cristi</i>	Religious treatise/Mysticism – Instruction (prose)	80,477 (137,288) (Including 372 words of <i>Meditatio de Passione Cristi</i>)
	<i>The Abbey of the Holy Ghost</i>	Sermon – Instruction, religious (prose)	4,571 (4,571)
Period II	<i>Alphabet of Tales Parts I and II</i>	Fiction/Narrative – Instruction, religious (prose)	180,913
	<i>Methodius, The Bygynnyng of the World</i>	Religious treatise – Instruction, religious (prose)	3,476 (3,476)
	Richard Rolle of <i>Hampole, Yorkshire Writers (rollhor1.rtf)</i>	Religious treatise/Mysticism – Instruction (prose)	19,716 (137,288)
	<i>The Mirror of St. Edmund</i>	Sermon – Instruction, religious (prose)	14,395 (14,395)
			305,961 (489,511)

¹¹⁷ The word count of a text after editing is given here, where relevant, in addition to word counts, which are adjoined in brackets, for each text provided by Markus (2006: §5.2). However, it is not clear whether the word count contains typographical practices and editorial comments in the *ICoMEP*, except for the following information: “[t]he ... list provides information about the exact sizes” (Markus 2006: §5).

Table 60*NME Texts from the TEAMS*

	Text	Type of text	Word count
Period II	<i>Octavian</i>	Romance (verse)	11,510
	<i>Sir Eglamour of Artois</i>	Romance (verse)	8,670
			20,180

Table 61

ESc Texts from the HCOS

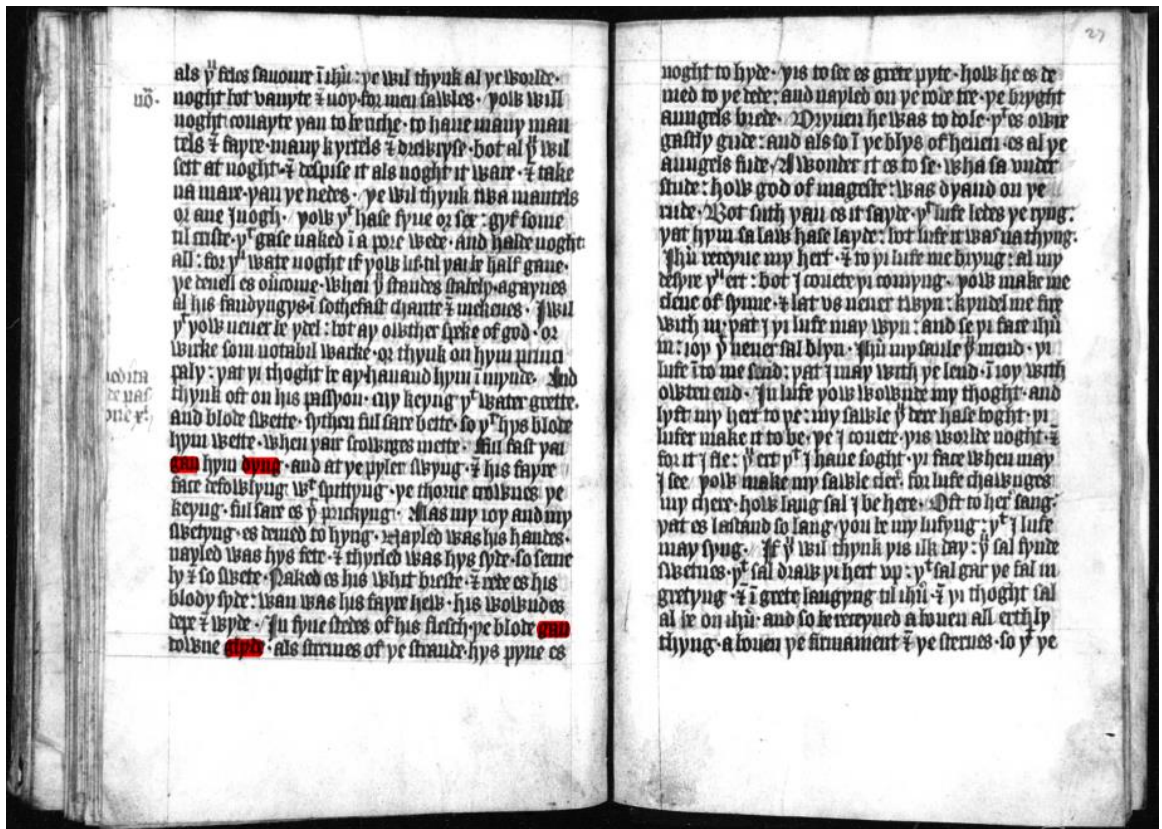
	Text	Type of text	Word count¹¹⁸
Period I	<i>The Douglas Book</i>	Letter Non-Private (prose)	1,396 (N/A)
	<i>Ratis Raving and Other Early Scots Poems on Morals: Craft of Deyng</i>	Education Treatise – Instruction, religious (prose)	3,243 (3,200)
	<i>Extracts from the Council Register of the Burgh of Aberdeen, 1398-1570</i>	Local Record – Stat (prose)	4,876 (4,900)
	<i>The Acts of Parliaments of Scotland</i>	Law – Stat (prose)	27,212 (27,300)
	<i>The Prose Works of Sir Gilbert Hay</i>	Education Treatise – Instruction, secular (prose)	28,745 (28,800)
Period II	<i>Charters and Documents Relating to the Burgh of Peebles</i>	Local Record – Stat (prose)	11,111 (N/A)
	<i>Ratis Raving and Other Early Scots Poems on Morals: The Vertewis of the Mess</i>	Education Treatise – Instruction, religious (prose)	823 (800)
	<i>Ratis Raving and Other Early Scots Poems on Morals: Dicta Salomonis</i>	Education Treatise – Instruction, religious (prose)	6,402 (6,400)
	<i>The Porteous of Noblenes</i>	Education Treatise – Instruction, secular (prose)	3,763 (N/A)
	<i>De Irelandia’s The Meroure of Wyssdome</i>	Education Treatise – Instruction, religious (prose)	8,329 (8,300)
			95,900

¹¹⁸ In a similar way to the *ICoMEP*, it is not clear whether Meurman-Solin (1995a) includes typographical practices and editorial comments in the word count in the *HCOS*. There is also a slight discrepancy between word counts given by Meurman-Solin (1995a: 58) and *ICAME CORPUS MANUALS*. While the former acknowledges this, word counts from both sources are provided here, with the latter being the main point of reference.

Table 62*ESc Texts from the TEAMS*

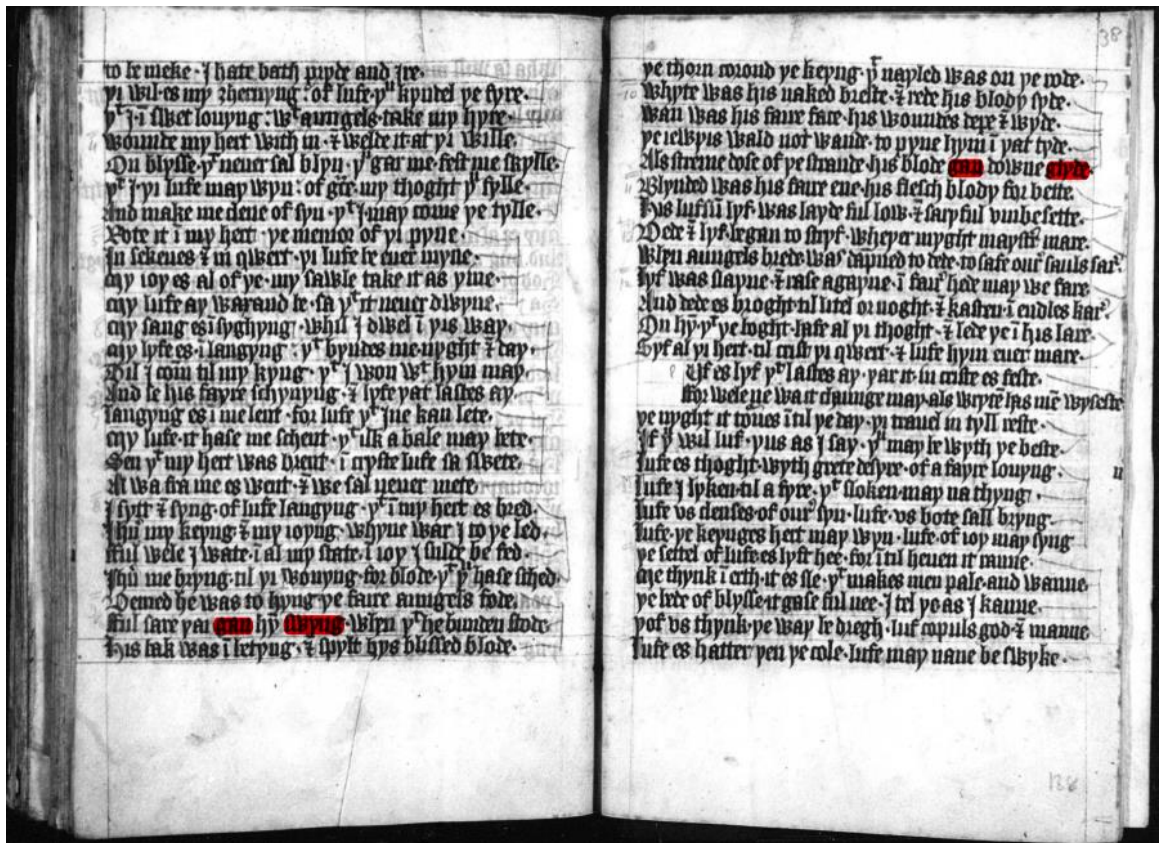
	Text	Type of text	Word count
Period I	<i>St. Andrew and the Three Questions</i>	Legend (verse)	1,833
	<i>The Life of St. Julian Hospitaller</i>	Legend (verse)	1,874
			3,707

Appendix Two



(Cambridge University MS DD.v.64 Part 3; emphasis mine)

Appendix Three



(Cambridge University MS DD.v.64 Part 3; emphasis mine)

Bibliography

- A Linguistic Atlas of Late Mediaeval English*. (1986). (A. McIntosh, Ed.) Aberdeen: Aberdeen University Press.
- A Linguistic Atlas of Older Scots*. (2008-2013). (K. Williamson, Producer, & The University of Edinburgh). Retrieved from <http://www.lel.ed.ac.uk/ihd/laos1/laos1.html>
- A Representative Corpus of Historical English Registers*. (D. Biber, & E. Finegan, Compilers). Manchester: Linguistics and English Language, University of Manchester.
- Aaron, J. (1993). *The Little, Brown Compact Handbook*. New York: Harper Collins College Publishers.
- Aberdeen Burgh Registers 1398-1511*. (n.d.). ScotlandsPlaces. Retrieved from <http://scotlandspplaces.gov.uk/digital-volumes/burgh-records/aberdeen-burgh-registers>
- Allen, R. S. (1989). *Richard Rolle: The English Writings*. New York, Mahwah: Paulist Press.
- An Electronic Version of A Linguistic Atlas of Late Mediaeval English*. (2013). (Benskin, M., Laing, M., Karaiskos, V., & K. Williamson, Compilers). Retrieved from <http://www.lel.ed.ac.uk/ihd/elalme/elalme.html>
- Andersen, H. (1973). Abductive and deductive change. *Language*, 49, 765–93.
- Antipova, Y. Y. (1963). Glago'noe zamescenie v drevne angliiskom yazyke. *Vestnik Leningradskogo Gosudarstvennogo Universiteta*. Series: Literature, History. *Language*, 8 (2), 125-36.
- Anttila, R. (1989). *Historical and Comparative Linguistics* (2nd ed.). Current Issues in Linguistic Theory, 6. Amsterdam: John Benjamins.
- Anttila, R. (2003). Analogy: The warp and woof of cognition. In B. D. Joseph, & R. D. Janda (Eds.), *The Handbook of Historical Linguistics* (pp. 425-440). Oxford: Blackwell.
- Ard, J. (1976). Rebracketing in diachronic syntax and Montague Grammar. In S. B. Steever, C. A. Walker, & S. S. Mufwene (Eds.), *Papers from the Parasession of Diachronic Syntax, April 22, 1976* (pp. 1-8). Chicago: Chicago Linguistic Society.

- Arnovick, L. K. (2012). Middle English: Literary language. In A. Berg, & L. J. Brinton (Eds.), *English Historical Linguistics: An International Handbook* (Vol. 1: 551-575). Berlin: Walter de Gruyter.
- Atkins, S., Clear, J., & Ostler, N. (1992). Corpus design criteria. *Literary and Linguistic Computing*, 7 (1), 1-16.
- Attridge, D. (2013). *Moving Words: Forms of English Poetry*. Oxford: Oxford University Press.
- Attridge, D. (1995). *Poetic Rhythm: An Introduction*. Cambridge: Cambridge University Press.
- Bach, E. (1986). The algebra of events. *Linguistics and Philosophy*, 9 (1), 5-16.
- Bache, C., & Davidsen-Nielsen, N. (1997). *Mastering English: An Advanced Grammar for Non-native and Native Speakers*. Berlin: Walter de Gruyter.
- Barnet, S., & Stubbs, M. (1995). *Barnet and Stubbs's Practical Guide to Writing*. New York: Harper Collins College Publishers.
- Bath, M., & Furniss, T. ([1996] 2007). *Reading Poetry: An Introduction* (2nd ed.). Harlow: Pearson Education.
- Beal, J. C., Corrigan, K. P., & Moisl, H. L. (2007). Taming digital voices and texts: Models and methods for handling unconventional synchronic corpora. In J. C. Beal, K. P. Corrigan, & H. L. Moisl (Eds), *Creating and Digitizing Language Corpora* (Vol. 2: 1-15). London: Palgrave Macmillan.
- Beschorner, F. (1920). Verbale Reime bei Chaucer. In *Studien zur Englischen Philologie*. Halle a. S.: Verlag von Max Niemeyer.
- Bhat, D. N. (1999). *The Prominence of Tense, Aspect, and Mood*. Studies in Language Companion Series, 49. Amsterdam: John Benjamins.
- Biber, D. (1988). *Variation Across Speech and Writing*. Cambridge: Cambridge University Press.
- Bible, King James Version: Hebrews*. (1997). (D. L. Services, Producer, & University of Michigan Library). Retrieved from M Library: Digital Collections <http://quod.lib.umich.edu/cgi/k/kjv/kjv-idx?type=DIV1&byte=5277078>
- Binnick, R. I. (1991). *Time and the Verb: A Guide to Tense & Aspect*. Oxford: Oxford University Press.

- Blake, N. F. (1977). *The English Language in Medieval Literature*. London: Dent.
- Blank, A. (1997). *Prinzipien des Lexikalischen Bedeutungswandels am Beispiel der Romanischen Sprachen*. Tübingen: Niemeyer.
- Blank, P. (1996). *Broken English: Dialects and the Politics of Language in Renaissance Writings*. London: Routledge.
- Boase-Beier, J. (2011). *A Critical Introduction to Translation Studies*. Continuum Critical Introductions to Linguistics. London: Bloomsbury.
- Bock, H. (1931). Studien zum präpositionalen Infinitiv und Akkusativ mit dem to-Infinitiv. *Anglia – Zeitschrift für englische Philologie*, 55, 114-176.
- Boogaart, R. (2004). Aspect and Aktionsart. In G. Booij, C. Lehmann, J. Mugdan, S. Skopeteas, & W. Kesselheim (Eds.), *Morphologie: Ein Internationales Handbuch Zur Flexion und Wortbildung* (Vol. 2: 1165-1180). Berlin: Walter de Gruyter.
- Bosworth, J. (2010). *Bosworth-Toller Anglo-Saxon Dictionary* (T. Northcote Toller, Editor, S. Christ, O. Tichý, Producers, & Faculty of Arts, Charles University in Prague). Retrieved from <http://www.bosworthtoller.com/>
- Boucher, J. (1832). *Boucher's Glossary of Archaic and Provincial Words: A Supplement to the Dictionaries of the English Language, Particularly Those of Dr. Johnson and Dr. Webster*. London: Printed for Black, Young, and Young.
- Breban, T. (2010). *English Adjectives of Comparison: Lexical and Grammaticalized Uses*. Topics in English Linguistics, 63. Berlin: De Gruyter Mouton.
- Brinton, L. (1981). *The Historical Development of Aspectual Periphrases in English* (Doctoral dissertation, University of California, Berkeley). Available from ProQuest Dissertations & Theses Global <http://search.proquest.com.ezproxy.is.ed.ac.uk/pqdtft/docview/303007215/B3FCD15377694DBFPQ/1?accountid=10673>
- Brinton, L. (1983). Criteria for distinguishing the non-aspectual functions of ME ginnen. *General Linguistics*, 23, 235-245.
- Brinton, L. (1988). *The Development of English Aspectual Systems: Aspectualizers and Post-verbal Particles*. Cambridge Studies in Linguistics, 49. Cambridge: Cambridge University Press.

- Brinton, L. (1996). *Pragmatic Markers in English: Grammaticalization and Discourse Functions*. Topics in English Linguistics, 19. Berlin: Mouton de Gruyter.
- Brinton, L., & Traugott, E. C. (2005). *Research Surveys in Linguistics: Lexicalization and Language Change*. Cambridge: Cambridge University Press.
- Brown, M. H. (2004). James I (1394-1437). *Oxford Dictionary of National Biography*. Retrieved from <http://www.oxforddnb.com/view/article/14587>
- Brown, W. H., Jr. (1970). *A Syntax of King Alfred's 'Pastoral Care'*. The Hague: Mouton.
- Bugaj, J. (2004). *Middle Scots Inflectional System in the South-West of Scotland: Studies in English Medieval Language and Literature*. Frankfurt am Main: Peter Lang.
- Burrow, J. A. (2008). *Medieval Writers and their Work*. Retrieved from EBook Central database
<https://ebookcentral.proquest.com/lib/ed/detail.action?docID=415532>
- Burrow, J. A., & Turville-Petre, T. ([1992] 1996). *A Book of Middle English* (2nd ed.). Oxford: Blackwell.
- Bybee, J. (1985). *Morphology. A Study of the Relations Between Meaning and Form*. Typological Studies in Language, 9. Amsterdam: John Benjamins.
- Bybee, J. (2003). Mechanisms of change in grammaticization: The role of frequency. In B. D. Joseph, & R. D. Janda (Eds.), *The Handbook of Historical Linguistics* (pp. 602-623). Oxford: Blackwell.
- Bybee, J. (2006). Where does grammar come from? In E. M. Rickerson, & B. Hilton (Eds.), *The Five-Minute Linguist: Bite-Sized Essays on Language and Languages* (pp. 50-53). Oakville: Equinox.
- Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: The University of Chicago Press.
- Callaway, M. (1913). *The Infinitive in Anglo-Saxon*. Washington: The Carnegie Institution of Washington.

- Cambridge University MS DD.v.64 Part 3. Microfilm. Cambridge: Cambridge University Press.
- Campbell, K. (1975). *The Seven Sages of Rome: Edited from the Manuscripts, with Introduction, Notes, and Glossary*. Genève: Slatkine Reprints.
- Campbell, L. (2001). What's wrong with grammaticalization? *Language Sciences*, 23, 113-161.
- Campbell, L. (2013). *Historical Linguistics: An Introduction* (3rd ed.). Edinburgh: Edinburgh University Press.
- Carlson, L. (1981). Aspect and quantification. In P. J. Tedeschi, & A. Zaenen, *Syntax and Semantics: Tense and Aspect* (Vol. 14: 31-64). New York: Academic Press.
- Chambers, E. K. (1947). *English Literature at the Close of the Middle Ages*. Oxford: Clarendon Press.
- Cole, M. (2014). *Old Northumbrian Verbal Morphosyntax and the (Northern) Subject Rule*. NOWELE Supplement Series, 25. Amsterdam: John Benjamins.
- Comrie, B. (1976). *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press.
- Corbett, G. G. (2013) Periphrasis and possible lexemes. In M. Chumakina, & G. G. Corbett (Eds.), *Periphrasis: The Role of Syntax and Morphology in Paradigms. Proceedings of the British Academy 180* (7) (pp. 169-189). Oxford: Oxford University Press for The British Academy.
- Corbett, J., McClure, J. D., & Stuart-Smith, J. (2003). A brief history of Scots. In J. Corbett, J. D. McClure, & J. Stuart-Smith (Eds.), *The Edinburgh Companion to Scots* (pp. 1-16). Edinburgh: Edinburgh University Press.
- Corpus of Middle English Prose and Verse*. (2006). (F. McSparran, Editor, & The Humanities Text Initiative, University of Michigan). Retrieved from <http://quod.lib.umich.edu/c/cme/>
- Craigie, W. A. (1931). *A Dictionary of the Older Scottish Tongue: From the Twelfth Century to the End of the Seventeenth* (Vol. 1: A-C). London: Humphrey Milford, Oxford University Press.

- Croft, W. (2000). *Explaining Language Change: An Evolutionary Approach*. Harlow: Pearson Education Limited.
- Crook, E. J. (1970). *Handbook of Middle English Grammar: Phonology*. (Unpublished PhD Dissertation). Urbana-Champaign: University of Illinois.
- Cuesta, J. F., & Silva, I. S. (2008). Does Old Northumbrian exist?: Northern varieties of Old English. In M. F. G-B. Giner, P. S. García, C. M. Granado, E. P. Iglesias, & J. A. J. Torresquesana (Eds.), *Multidisciplinary Studies in Language and Literature: English, American and Canadian: In Memoriam Gudelia Rodríguez Sánchez* (pp. 49-58). Salamanca: Ediciones Universidad de Salamanca.
- Dahl, Ö. (1985). *Tense and Aspect Systems*. Oxford: Blackwell.
- Dalrymple, R. (2004). Genre. In R. Dalrymple (Ed.), *Middle English Literature: A Guide to Criticism* (pp. 55-58). Oxford: Blackwell.
- Damian-Grint, P. (1999). *The New Historians of the Twelfth-Century Renaissance: Inventing Vernacular Authority*. Woodbridge: The Boydell Press.
- De la Cruz, J. M. (1975). Old English pure prefixes: Structure and function. *Linguistics*, 145, 47-79.
- De Saussure, F. (1962). *Cours de Linguistique Générale*. Paris: Payot.
- De Smet, H. (2013). *Spreading Patterns: Diffusional Change in the English System of Complementation*. Oxford Studies in the History of English. Oxford: Oxford University Press.
- Declerck, R., Reed, S., & Cappelle, B. (2006). *The Grammar of the English Verb Phrase: Volume 1: The Grammar of the English Tense System*. Topics in English Linguistics 60-1. (B. Kortmann, & E. C. Traugott, Eds.) Berlin: Mouton de Gruyter.
- Denison, D. (1985). The origins of periphrastic Do: Ellegård and Visser reconsidered. In R. Eaton, O. Fischer, W. F. Koopman, & F. van der Leek (Eds.), *Papers from the 4th International Conference on English Historical Linguistics, Amsterdam, April 10–13, 1985*. Current Issues in Linguistic Theory, 41 (pp. 45-60). Amsterdam: John Benjamins.
- Denison, D. (1993). *English Historical Syntax*. London: Longman.

- Dér, C. (2013). Grammaticalization: A specific type of semantic, categorical, and prosodic change. *Berliner Beiträge zur Hungarologie*, 18, 160-179. Retrieved from https://www.slawistik.hu-berlin.de/de/fachgebiete/ungarlit/publ/BBH/bbh18b/bbh18_der.pdf
- Deutschbein, M. (1917). *System der Neuenglischen Syntax*. Cöthen: Otto Schulze.
- Dictionary of Old English Corpus in Electronic Form*. (2005). (A. di Paolo Healey, Editor, & University of Oxford Text Archive) Available from <http://ota.ox.ac.uk/desc/2488>
- Dictionary of Old English: A to G Online*. (2007). (A. Cameron, A. Crandell Amos, A. di Paolo Healey, Editors, & Dictionary of Old English Project). Available from <http://tapor.library.utoronto.ca/doe/dict/index.html>
- Dictionary of the Scots Language*. (2004). (S. Rennie, Editor, & Scottish Language Dictionaries). Retrieved from www.dsl.ac.uk
- Diewald, G. (1997). *Grammatikalisierung*. Tübingen: Niemeyer.
- Diewald, G. (2002). A model for relevant types of contexts in grammaticalization. In I. Wischer, & G. Diewald (Eds.), *New Reflections on Grammaticalization*. *Typological Studies in Language*, 49 (pp. 103-120). Amsterdam: John Benjamins.
- Diewald, G. (2010). On some problem areas in grammaticalization studies. In K. Stathi, E. Gehweiler, & E. König (Eds.), *Grammaticalization: Current Views and Issues*. *Studies in Language Companion Series*, 119 (pp. 17-50). Amsterdam: John Benjamins.
- Diewald, G., & Smirnova, E. (2012). "Paradigmatic integration": The fourth stage in an expanded grammaticalization scenario. In D. Kristin, B. Tine, B. Lieselotte, & M. Tanja (Eds.), *Grammaticalization and Language Change: New Reflections*. *Studies in Language Companion Series*, 130 (pp. 111-134). Amsterdam: John Benjamins.
- Dik, S. C. (1977). Inductive generalizations in semantic change. In P. J. Hopper (Ed.), *Lehmann, Studies in Descriptive and Historical Linguistics: Festschrift for Winfred P. Lehmann*. *Current Issues in Linguistic Theory*, 4 (pp. 283-300). Amsterdam: John Benjamins.

- Dorr, B. J. (1993). *Machine Translation: View from the Lexicon*. Cambridge, Massachusetts: The MIT Press.
- Douthwaite, J. (2000). *Towards a Linguistic Theory of Foregrounding*. Alessandria: Edizioni dell'Orso.
- Dowty, D. R. (1979). *Word Meaning and Montague Grammar: The Semantics of Verbs and Rimes in Generative Semantics and in Montague's PTQ*. Dordrecht: Reidel.
- Dry, H. A. (1992). Foregrounding: An assessment. In S. J. J. Hwang, & W. R. Merrifield (Eds.). *Language in Context: Essays for Robert E. Longacre*. Summer Institute of Linguistics Publications in Linguistics, 107 (pp. 435-450). The Summer Institute of Linguistics and The University of Texas at Arlington.
- Duffell, M. J. (2008). *A New History of the English Metre*. Studies in Linguistics, 5. Oxford: Legenda.
- Duncan, A. A. (2004). Barbour, John (c. 1330-1395). *Oxford Dictionary of National Biography*. Available from <http://www.oxforddnb.com/view/article/1336>
- Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19 (1), 61-74.
- Dupré, L. (1995). *Bugs in Writing: Aids to Debugging your Prose*. Reading, Massachusetts: Addison-Wesley.
- Earle, J. (1887). *The Philology of the English Tongue* (3rd ed.). Oxford: The Clarendon Press.
- Eckardt, R. (2006). *Meaning Change in Grammaticalization: An Enquiry into Semantic Reanalysis*. Oxford: Oxford University Press.
- Einenkel, E. (1891). Der Infinitiv im Mittelenglischen. *Anglia*, 13, 79-104.
- Ellegård, A. (1953). *The Auxiliary Do: The Establishment and Regulation of its Use in English*. Göteborg: Elanders Boktryckeri Aktiebolag.
- Erman, B. (1992). Female and male usage of pragmatic expressions in same-sex and mixed-sex interaction. *Language Variation and Change*, 4, 217-234.
- Eyre-Todd, G. ([1907] 1996). Preface. *John Barbour: The Bruce: The History of Robert the Bruce King of Scots*. (G. Eyre-Todd, Trans.). Edinburgh: Mercat Press.

- Fischer, O. (1992). Syntax. In N. Blake (Ed.), *The Cambridge History of the English Language* (Vol. 2: 207-408). Cambridge: Cambridge University Press.
- Fischer, O. (1994). The development of quasi-auxiliaries in English and changes in word order. *Neophilologus*, 78, 137-164.
- Fischer, O. (2007). *Morphosyntactic Change: Functional and Formal Perspectives*. Oxford: Oxford University Press.
- Fischer, O., & Rosenbach, A. (2000). Introduction. In O. Fischer, A. Rosenbach, & D. Stein (Eds.), *Pathways of Change: Grammaticalization in English*. Studies in Language Companion Series, 53 (pp. 1-38). Amsterdam: John Benjamins.
- Fischer, O., & van der Wurff, W. (2006). Syntax. In R. Hogg, & D. Denison (Eds.), *A History of the English Language*. Cambridge: Cambridge University Press.
- Fischer, O., van Kemenade, A., Koopman, W., & van der Wurff, W. (2000). *The Syntax of Early English*. Cambridge Syntax Guides. Cambridge: Cambridge University Press.
- Fisiak, J. (1996). *A Short Grammar of Middle English: Orthography, Phonology and Morphology*. Warszawa: Wydawnictwo Naukowe PWN.
- Fleischman, S. (1985). Discourse functions of tense-aspect oppositions in narrative: Toward a theory of grounding. *Linguistics*, 23, 851-881.
- Fleischman, S. (1990). *Tense and Narrativity: From Medieval Performance to Modern Fiction*. Austin: University of Texas Press.
- Francis, N. W. (1979). Problems of assembling and computerizing large corpora. In H. Bergenholtz, & B. Schaefer (Eds.), *Empirische Textwissenschaft. Aufbau und Auswertung von Text-Corpora* (pp. 110-123). Königstein im Taunus: Scriptor.
- Francis, N. W. (1992). Language corpora B.C. In J. Svartvik (Ed.), *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82*. Stockholm, 4-8 August 1991. Trends in Linguistics. Studies and Monographs (pp. 17-32). Berlin: Mouton de Gruyter.
- Frank, W. C. (1912). The Wakefield group in Towneley. *The Journal of English and Germanic Philology*, 11 (2), 244-262.
- French, W. H., & Hale, C. B. (Eds.). (1930). *Middle English Metrical Romances*. New York: Prentice-Hall.

- Fulk, R. D. (2012). *An Introduction to Middle English: Grammar and Texts*. Peterborough, Ontario: Broadway Press.
- Funke, V. (1922). Die Fügung ginnen mit dem Infinitiv im Mittelenglischen. *Englische Studien*, 56, 1-27.
- Gaeta, L. (1998). Some remarks on analogy, reanalysis and grammaticalization. In A. G. Ramat, & P. J. Hopper (Eds.), *The Limits of Grammaticalization*. *Typological Studies in Language*, 37 (pp. 89-105). Amsterdam: John Benjamins.
- Gardela, W. (2014). Morpho-syntactic variation of the present participle in the late 14th and the 15th century Northern English and Scots texts. *English Studies*, 95 (2), 149-176.
- Garvin, P. (1964). The standard language problem: Concepts and methods. In D. Hymes (Ed.), *Language in Culture and Society: A Reader in Linguistics and Anthropology* (pp. 521-526). New York: Harper & Row.
- Geeraerts, D. (1997). *Diachronic Prototype Semantics: A Contribution to Historical Lexicology*. Oxford: Clarendon Press.
- Gimson, A. C. (1980). *An Introduction to the Pronunciation of English*. London: Arnold.
- Gisborne, N. (2007). Dynamic modality. *Skase Journal of Theoretical Linguistics*, 4 (2), 44-61. Retrieved from <http://www.skase.sk/Volumes/JTL09/>
- Givón, T. (1979). *On Understanding Grammar*. London: Academic Press.
- Givón, T. (1995). *Functionalism and Grammar*. Amsterdam: John Benjamins.
- Görlach, M. (2002). *A Textual History of Scots*. Heidelberg: Winter.
- Gramley, S. (2012). *The History of English: Introduction*. London: Routledge.
- Greenbaum, S., & Quirk, R. (1990). *A Student's Grammar of the English Language*. Harlow: Longman.
- Haiman, J. (1977). Reinterpretation. *Language*, 53 (2), 312-328.
- Haiman, J. (1991). From V/2 to subject clitics: Northern Italian. In E. C. Traugott, & B. Heine (Eds.), *Approaches to Grammaticalization*. *Typological Studies in Language* 19: 2 (pp. 135-157). Amsterdam: John Benjamins.

- Haiman, J. (1994). Ritualization and the development of language. In W. Pagliuca (Ed.), *Perspectives on Grammaticalization*. Current Issues in Linguistic Theory, 109 (pp. 3-28). Amsterdam: John Benjamins.
- Halliday, M. A. (1961). Categories of the theory of grammar. *Word*, 17, 241-292.
- Halliday, M. A. (1970). Language structure and language function. In J. Lyons (Ed.), *New Horizons in Linguistics* (pp. 140-165). Harmondsworth: Penguin.
- Halliday, M. A. (1979). Modes of meaning and modes of expression: Types of grammatical structure, and their determination by different semantic functions. In D. Allerton, E. Carney, & D. Holdcroft (Eds.), *Function and Context in Linguistic Analysis: A Festschrift for William Haas* (pp. 57-79). Cambridge: Cambridge University Press.
- Harding, D. C. W. ([1976] 2010). *Words into Rhythm: English Speech Rhythm in Verse and Prose*. Cambridge: Cambridge University Press.
- Harris, A. C., & Campbell, L. (1995). *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.
- Haspelmath, M. (1998). Does grammaticalization need reanalysis? *Studies in Language*, 22 (2), 315-351.
- Haspelmath, M. (1999). Why is grammaticalization irreversible? *Linguistics*, 37 (6), 1043-1068.
- Hatav, G. (1993). The aspect system in English: An attempt at a unified analysis. *Linguistics*, 31 (2), 209-237.
- Häusermann, H. W. (1930). *Studien zu den Aktiosarten im Frühmittelenglischen*. Wien und Leipzig: Wilhelm Braumüller: Universitäts-Verlagsbuchhandlung.
- Head, G. (1997). *Studies in the Language, Palaeography and Codicology of MS Edinburgh, National Library of Scotland, Advocates' 19.2.2* (Doctoral dissertation, University of Glasgow). Retrieved from Enlighten: Theses <http://theses.gla.ac.uk/2989/>
- Heine, B. (1992). Grammaticalization chains. *Studies in Language*, 16 (2), 335-368.
- Heine, B. (1993). *Auxiliaries: Cognitive Forces and Grammaticalization*. Oxford: Oxford University Press.
- Heine, B. (1997). *Possession: Cognitive Sources, Forces, and Grammaticalization*. Cambridge: Cambridge University Press.

- Heine, B., & Kuteva, T. (2002). *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Heine, B., & Kuteva, T. (2005). *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.
- Heine, B., & Reh, M. (1984). *Grammaticalization and Reanalysis in African Languages*. Hamburg: Buske Helmet Verlag.
- Heine, B., & Reh, M. (1994). *Patterns of Grammaticalization in African Languages. AKUP 47 (Arbeiten des Kölner Universalien-Projekts)*. Institut für Sprachwissenschaft. Cologne: University of Cologne.
- Heine, B., Claudi, U., & Hünnemeyer, F. (1991a). From cognition to grammar: Evidence from African languages. In E. C. Traugott, & B. Heine (Eds.), *Approaches to Grammaticalization*. Typological Studies in Language 19: 1 (pp. 149-188). Amsterdam: John Benjamins.
- Heine, B., Claudi, U., & Hünnemeyer, F. (1991b). *Grammaticalization: Conceptual Framework*. Chicago: University of Chicago Press.
- Hiatt, A. (2007). Genre without system. In P. Strohm (Ed.), *Oxford Twenty-First Century Approaches to Literature: Middle English* (pp. 277-294). Oxford: Oxford University Press.
- Hiltunen, R. (1983). *The Decline of the Prefixes and the Beginnings of the English Phrasal Verb*. Turku: Turun Yliopisto.
- Hobsbaum, P. (1996). *Metre, Rhythm and Verse Form*. London: Routledge.
- Hock, H. H. (1991). *Principles of Historical Linguistics* (2nd ed.). Berlin: Mouton de Gruyter.
- Hoffmann, S. (2004). Are low-frequency complex prepositions grammaticalized? In H. Lindquist, & C. Mair (Eds.), *Corpus Approaches to Grammaticalization in English*. Studies in Corpus Linguistics, 13 (pp. 171-210). Amsterdam: John Benjamins.
- Hogg, R. M. (2002). *An Introduction to Old English*. Edinburgh: Edinburgh University Press.
- Hogg, R. M., & Fulk, R. D. (2011). *A Grammar of Old English, Volume 2: Morphology*. Oxford: Wiley-Blackwell.

- Homann, E. R. (1954). Chaucer's use of 'gan'. *The Journal of English and Germanic Philology*, 53 (3), 389-398.
- Hopper, P. (1979a). Some observations on the typology of focus and aspect in narrative language. *Studies in Language*, 3 (1), 37-64.
- Hopper, P. (1979b). Aspects and foregrounding in discourse. In T. Givón (Ed.), *Discourse and Syntax: Syntax and Semantics 2* (pp. 213-241). New York: Academic Press.
- Hopper, P. (1991). On some principles of grammaticalization. In E. C. Traugott, & B. Heine (Eds.), *Approaches to Grammaticalization*. Typological Studies in Language 19: 1 (pp. 17-35). Amsterdam: John Benjamins.
- Hopper, P. (1992). Grammaticalization. In W. Bright (Ed.), *International Encyclopedia of Linguistics* (Vol. 2: 79-81). Oxford: Oxford University Press.
- Hopper, P. (1998). The paradigm at the end of the universe. In A. Giacalone-Ramat, & P. Hopper (Eds.), *The Limits of Grammaticalization*. Typological Studies in Language, 37 (pp. 147-158). Amsterdam: John Benjamins.
- Hopper, P., & Thompson, S. (1980). Transitivity in grammar and discourse. *Language*, 56, 251-299.
- Hopper, P., & Traugott, E. C. ([1993] 2003). *Grammaticalization* (2nd ed.). Cambridge: Cambridge University Press.
- Horobin, S., & Smith, J. (2002). *An Introduction to Middle English*. Oxford: Oxford University Press.
- Horstman, C. (Ed.). (1895). *Yorkshire Writers: Richard Rolle of Hampole an English Father of the Church and his Followers* (Vol. 1). London: Swan Sonnenschein & Co.
- Horstman, C. (Ed.). (1896). *Yorkshire Writers: Richard Rolle of Hampole and his Followers* (Vol. 2). London: Swan Sonnenschein & Co.
- Horstmann, C. (Ed.). (1881-1882). *Barbour's des schottischen Nationaldichters Legendensammlung: nebst den Fragmenten seines Trojanerkrieges*. Heilbronn: Gebr. Henninger.
- Hoye, L. (1997). *Adverbs and Modality in English*. London: Longman.

- Huddleston, R., & Pullum, G. K. (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Huddleston, R., & Pullum, G. K. (2005). *A Student's Introduction to English Grammar*. Cambridge: Cambridge University Press.
- Hudson, S. (Ed.). (2006a). *Sir Eglamour of Artois: Introduction. Four Middle English Romances: Sir Isumbras, Octavian, Sir Eglamour or Artois, Sir Tryamour*. (University of Rochester). Retrieved from TEAMS: Middle English Texts Series
<http://d.lib.rochester.edu/teams/text/hudson-sir-eglamour-of-artois-introduction>
- Hudson, S. (Ed.). (2006b). *Octavian: Introduction. Four Middle English Romances: Sir Isumbras, Octavian, Sir Eglamour or Artois, Sir Tryamour*. (University of Rochester). Retrieved from TEAMS: Middle English Texts Series
<http://d.lib.rochester.edu/teams/text/hudson-octavian-introduction>
- Indurkha, B. (1992). *Metaphor and Cognition: An Interactionist Approach*. London: Kluwer Academic.
- Inoue, N., & Stokes, M. (2009). The caesura and the rhythmic shape of the A-verse in the poems of the alliterative revival. In A. Hall (Ed.), *Leeds Studies in English: New Series XL*, 1-26. Retrieved from
<http://digital.library.leeds.ac.uk/12227/1/LSE2009.pdf>
- Jamieson, J. (1808). *An Etymological Dictionary of the Scottish Language: Illustrating the Words in their Different Significations, by Examples from Ancient and Modern Writers: Showing their Affinity to Those of Other Languages, and Especially the Northern: Explaining Many Terms, which, Though Now Obsolete in England, Were Formerly Common to Both Countries: And Elucidating National Rites, Customs, and Institutions, in Their Analogy to Those of Other Nations: To Which is Prefixed* (Vol. 1). Edinburgh: Printed at the University Press, for W. Creech, A. Constable & Co., and W. Blackwood, Edinburgh, Longman, Hurst, Rees, & Orme, T. Cadell & W. Davies, and H. D. Symonds, London.
- Jeffries, L., & McIntyre, D. (2010). *Stylistics*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

- Johnson, E. L. (1993). *An "Alphabet of Tales": The Genre, Background, Date, and Provenance of the Text, with an Annotated Glossary* (Vol. 1 and 2). (Doctoral dissertation, University of Michigan, Michigan). Available from ProQuest Dissertations & Theses Global
<http://search.proquest.com/docview/304056425?accountid=10673>
- Jones, M. C., & Singh, I. (2005). *Exploring Language Change*. London: Routledge.
- Jordan, R. (1974). *Handbook of Middle English Grammar: Phonology*. (E. J. Crook, Trans.). The Hague: Mouton.
- Jucker, A. H. (2002). Discourse markers in Early Modern English. In R. Watts, & P. Trudgill (Eds.), *Alternative Histories of English* (pp. 210-230). London: Routledge.
- Jucker, A. H., & Ziv, Y. (1998). Introduction. In A. H. Jucker, & Y. Ziv (Eds.), *Discourse Markers: Descriptions and Theory*. Pragmatics & Beyond New Series, 57 (pp. 1-12). Amsterdam: John Benjamins.
- Jumpertz-Schwab, C. (1998). *The Development of the Scots Lexicon and Syntax in the 16th Century under the Influence of Translations from Latin*. Frankfurt am Main: Peter Lang.
- Kabakčiev, K. J. (2000). *Aspect in English: A "Common Sense" View of the Interplay Between Verbal and Nominal Referents*. London: Kluwer Academic Publishers.
- Kato, T. (1974). *A Concordance to the Works of Sir Thomas Malory*. Tokyo: University of Tokyo Press.
- Kenny, A. (1963). *Action, Emotion and Will*. London: Routledge.
- Kenstowicz, M. J. (1994). *Phonology in Generative Grammar*. Oxford: Blackwell.
- Kerkhof, J. (1966). *Studies in the Language of Geoffrey Chaucer*. Leiden: Universitaire Pers Leiden.
- King, A. (1997). The inflectional morphology of Older Scots. In C. Jones (Ed.), *The Edinburgh History of the Scots Language* (pp. 156-181). Edinburgh: Edinburgh University Press.
- Kiparsky, P. (1968). Linguistic universals and linguistic change. In E. W. Bach, & R. T. Harms (Eds.), *Universals in Linguistic Theory* (pp. 170-202). New York: Holt, Rinehart and Winston.

- Kiparsky, P. (1992). Analogy. In W. Bright (Ed.), *International Encyclopedia of Linguistics* (Vol. 1: 56-61). Oxford: Oxford University Press.
- Knoppers, G. N. (2013). *Jews and Samaritans: The Origins and History of Their Early Relations*. Oxford: Oxford University Press.
- Kopaczyk, J. (2013). Rethinking the traditional periodisation of the Scots language. In J. Cruickshank, & R. McColl (Eds.), *Forum for Research on the Languages of Scotland and Ireland: After the Storm: Papers from the Forum for Research on the Languages of Scotland and Ulster Triennial Meeting, Aberdeen 2012*. Retrieved from <https://www.abdn.ac.uk/pfrlsu/documents/Kopaczyk,%20Rethinking%20the%20traditional%20periodisation.pdf>
- Koziol, H. (1932). *Wiener Beiträge zur Englischen Philologie: Grundzüge der Syntax der Mittelenglischen Stabreimdichtungen*. Wien und Leipzig: Wilhelm Braumüller.
- Koziol, H. (1936). Die Entstehung der Umschreibung mit to do. *Germanische-Romanische Monatsschrift*, 24, 460-466.
- Kroch, A. (1989). Function and grammar in the history of English: Periphrastic do. In R. W. Fasold, & D. Schiffrin (Eds.), *Language Change and Variation. Current Issues in Linguistic Theory*, 52 (pp. 133-172). Amsterdam: John Benjamins.
- Krug, M. (1998). String frequency: A cognitive motivating factor in coalescence, language processing, and linguistic change. *Journal of English Linguistics*, 26 (4), 286-320.
- Kuryłowicz, J. (1964). *The Inflectional Categories of Indo-European*. Heidelberg: Winter.
- Kuryłowicz, J. (1965). The evolution of grammatical categories. *Diogenes*, 13 (51), 55-71.
- Kuteva, T. (2001). *Auxiliation: An Enquiry into the Nature of Grammaticalization*. Oxford: Oxford University Press.
- Kytö, M. (1996). *Manual to the Diachronic Part of the Helsinki Corpus of English Texts: Coding Conventions and Lists of Source Texts* (Department of English, University of Helsinki). Retrieved from <http://icame.uib.no/hc/>

- Lakoff, G., & Johnson, M. (1980). The metaphorical structure of the human conceptual system. *Cognitive Science*, 4 (2), 195-208.
- Langacker, R. W. (1977). Syntactic reanalysis. In C. N. Li (Ed.), *Mechanisms of Syntactic Change* (pp. 57-139). Austin: University of Texas Press.
- Lass, R. (1992). Phonology and morphology. In N. Blake (Ed.), *The Cambridge History of the English Language* (pp. 23-155). Cambridge: Cambridge University Press.
- Lass, R. (2006). Phonology and morphology. In R. Hogg, & D. Denison (Eds.), *A History of the English Language* (pp. 43-108). Cambridge: Cambridge University Press.
- Leech, G. (1971). *Meaning and the English Verb*. London: Longman.
- Leech, G. (1991). The state of the art in corpus linguistics. In K. Aijmer, & B. Altenberg (Eds.), *English Corpus Linguistics: Studies in Honour of Jan Svartvik* (pp. 8-29). London: Longman.
- Leech, G. (1992). Corpora and theories of linguistic performance. In J. Svartvik (Ed.), *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82. Stockholm, 4-8 August 1991*. Trends in Linguistics. Studies and Monographs (pp. 105-126). Berlin: Mouton de Gruyter.
- Leech, G. ([1969] 2013). *A Linguistic Guide to English Poetry*. Abingdon: Routledge.
- Leech, G. (2013). *Language in Literature: Style and Foregrounding*. London: Routledge.
- Leech, G., Hundt, M., Mair, C., & Smith, N. (2009). *Change in Contemporary English: A Grammatical Study*. Studies in English Language. Cambridge: Cambridge University Press.
- Lehmann, C. (1985). Grammaticalization: Synchronic variation and diachronic change. *Lingua e Stile* (20), 303-318.
- Lehmann, C. (1995). *Thoughts on Grammaticalization*. Munich: Lincom Europa.
- Lehmann, C. (2004). Theory and method in grammaticalization. In V. Ágel, F. H. Linke, A. Lüdeling, & A. T. Doris (Eds.), *Zeitschrift für germanistische Linguistik: Deutsche Sprache in Gegenwart und Geschichte* 32: 2 (pp. 152-187). Berlin: Mouton de Gruyter.

- Levinson, S. C. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge, Massachusetts: The MIT Press.
- Lewis, R. E. (Ed.). (2007). *Middle English Dictionary: Plan and Bibliography*. Ann Arbor: The University of Michigan Press.
- Lichtenberk, F. (1991). On the gradualness of grammaticalization. In E. C. Traugott, & B. Heine (Eds.), *Approaches to Grammaticalization*. Typological Studies in Language, 19:1 (37-80). Amsterdam: John Benjamins.
- Lightfoot, D. (1979). *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.
- Lightfoot, D. (1991). *How to Set Parameters: Arguments from Language Change*. Cambridge, Massachusetts: The MIT Press.
- Lightfoot, D. (2006). *How New Languages Emerge*. Cambridge: Cambridge University Press.
- Lloyd, A. L. (1979). *The Gothic Verb as a Model for a Unified Theory of Aspect, Actional Types, and Verbal Velocity*. Studies in Language Companion Series, 4. Amsterdam: John Benjamins.
- Lord, C. (1976). Evidence for syntactic reanalysis: From verb to complementizer in Kwa. In S. Steever, C. Walker, & S. Mufwene (Eds.), *Papers from the Parasession on Diachronic Syntax, April 22, 1976* (pp. 179-191). Chicago: Chicago Linguistic Society.
- Los, B. (1998). The rise of the *to*-infinitive as verb complement. *English Language and Linguistics*, 2 (1), 1-36.
- Los, B. (2000). *Onginnan/beginnan* with bare and *to*-infinitive in Ælfric. In O. Fischer, A. Rosenbach, & D. Stein (Eds.), *Pathways of Change: Grammaticalization in English*. Studies in Language Companion Series, 53 (pp. 521-274). Amsterdam: John Benjamins.
- Los, B. (2005). *Rise of the To-Infinitive*. Oxford: Oxford University Press.
- Los, B., Blom, C., Booij, G., Elenbaas, M., & van Kemenade, A. (2012). *Morphosyntactic Change. A Comparative Study of Particles and Prefixes*. Cambridge Studies in Linguistics, 134. Cambridge: Cambridge University Press.

- Lynch, M. (2007). *The Oxford Companion to Scottish History*. Oxford: Oxford University Press.
- Macafee, C. (n.d.). History of Scots to 1700. In *Dictionary of the Scots Language*. Retrieved from <http://www.dsl.ac.uk/about-scots/history-of-scots/>
- Macafee, C. (1992/93). A short grammar of Older Scots. *Scottish Language*, 11/12, 10-36.
- Madden, F. (1847). *Layamons Brut, or Chronicle of Britain; A Poetical Semi-Saxon Paraphrase of the Brute of Wace: Now First Published from the Cottonian Manuscripts in the British Museum: Accompanied by a Literal Translation, Notes, and a Grammatical Glossary*. London: The Society of Antiquaries of London.
- Mair, C. (2004). Corpus linguistics and grammaticalization theory: Statistics, frequencies, and beyond. In H. Lindquist, & C. Mair (Eds.), *Corpus Approaches to Grammaticalization in English*. Studies in Corpus Linguistics, 13 (pp. 121-150). Amsterdam: John Benjamins.
- Marchand, H. (1938/39). Syntaktische Homonymie. Das Umschreibende *do*. *Englische Studien*, 73, 227-252.
- Marchand, H. (1939). Review of Engblom 1938. *English Studies*, 21, 121-125.
- Markus, M. (1997). ME *can* and *gan* in context. In R. Hickey, & S. Puppel (Eds.), *Language History and Linguistic Modelling: A Festschrift for Jacek Fisiak on his 60th Birthday* (Vol. 1: 343-356). Berlin: Mouton de Gruyter.
- Markus, M. (2006). *Manual of the Innsbruck Prose Corpus: Version 2.2*. Innsbruck, Austria.
- Matisoff, J. A. (1991). Areal and universal dimensions of grammaticalization in Lahu. In E. C. Traugott, & B. Heine (Eds.), *Approaches to Grammaticalization*. Typological Studies in Language 19: 2 (pp. 383-453). Amsterdam: John Benjamins.
- Matthews, H. (1962). *Character & Symbol in Shakespeare's Plays: A Study of Certain Christian and Pre-Christian Elements in their Structure and Imagery*. Cambridge: Cambridge University Press.
- Mayor, J. B. ([1901] 2014). *Chapters on English Metre* (2nd ed.). Cambridge: Cambridge University Press.

- McClure, J. D. (1994). English in Scotland. In R. Burchfield (Ed.), *The Cambridge History of the English Language: English in Britain and Overseas. Origins and Developments* (Vol. 5: 23-93). Cambridge: Cambridge University Press.
- McEnery, T., & Hardie, A. (2012). *Corpus Linguistics: Method, Theory and Practice*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- McEnery, T., & Wilson, A. (2001). *Corpus Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.
- McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-Based Language Studies: An Advanced Resource Book*. Routledge Applied Linguistics. London: Routledge.
- McMahon, A. M. S. (1994). *Understanding Language Change*. Cambridge: Cambridge University Press.
- Meillet, A. (1912). L'évolution des formes grammaticales. *Scientia (Rivista di scienza)*, XII (XXVI, 6).
- Meillet, A. ([1912] 1958). *Linguistique Historique et Linguistique Generale*. Paris: Librairie Honore Champion.
- Mellor, M. D. (1995). *Aspects of Aspectual Verbs in English and Russian* (Doctoral dissertation, University of Edinburgh, Edinburgh). Retrieved from Edinburgh Research Archive
https://www.era.lib.ed.ac.uk/bitstream/handle/1842/525/Mellor_pdf?sequence=3&isAllowed=y
- Merriam-Webster's Encyclopedia of Literature*. (1995). Springfield, Massachusetts: Merriam-Webster.
- Metcalf, W. M. (Ed.). (1896). *Legends of the Saints in the Scottish Dialect of the Fourteenth Century: Edited from the Unique Manuscript in the University Library, Cambridge, with Introduction, Notes, and Glossarial Index* (Vol. 1). Edinburgh: William Blackwood and Sons for the Scottish Text Society.
- Meurman-Solin, A. (1992). On the morphology of verbs in Middle Scots: Present and present perfect indicative. In M. Rissanen, O. Ihalainen, T. Nevalainen, & I. Taavitsainen (Eds.), *History of Englishes: New Methods and*

- Interpretations in Historical Linguistics*. Topics in English Linguistics, 10 (pp. 611-623). Berlin: Mouton de Gruyter.
- Meurman-Solin, A. (1993a). Older Scots. In M. Rissanen, M. Kytö, & M. Palander-Collin (Eds.), *Early English in the Computer Age: Explorations through the Helsinki Corpus* (pp. 75-82). Berlin: Mouton de Gruyter.
- Meurman-Solin, A. (1993b). Periphrastic *do* in sixteenth- and seventeenth-century Scots. In M. Rissanen, M. Kytö, & M. Palander-Collin (Eds.), *Early English in the Computer Age: Explorations through the Helsinki Corpus* (pp. 235-251). Berlin: Mouton de Gruyter.
- Meurman-Solin, A. (1993c). *Variation and Change in Early Scots Prose*. Annales Academiae Scientiarum Fennicae. Dissertationes Humanarum Litterarum ed. Helsinki: Suomalainen Tiedeakatemia.
- Meurman-Solin, A. (1995). A new tool: The Helsinki Corpus of Older Scots (1450-1700). *ICAME Journal*, 19 (4), 49-62.
- Middle English Dictionary*. (2001). (F. McSparran, Editor, & The Regents of the University of Michigan). Available from <http://quod.lib.umich.edu/m/med/>
- Minkova, D. (2003). *Alliteration and Sound Change in Early English*. Cambridge: Cambridge University Press.
- Minkova, D. (2007). The forms of verse. In P. Brown (Ed.), *A Companion to Medieval English Literature and Culture* (pp. 176-197). Oxford: Blackwell.
- Minkova, D. (2012). Prosody. In A. Bergs, & Laurel J. Brinton (Eds.), *English Historical Linguistics. An International Handbook*. Handbooks of Linguistics and Communication Science, 34: 1 (pp. 113-129). Berlin: Mouton de Gruyter.
- Moder, C. L. (1992). Rules and analogy. In G. W. Davies, & G. K. Iverson (Eds.), *Explanation in Historical Linguistics*. Current Issues in Linguistic Theory, 84 (pp. 179-192). Amsterdam: John Benjamins.
- Moessner, L. (1989). *Early Middle English Syntax*. Tübingen: Max Niemeyer Verlag.
- Moessner, L. (1997). The syntax of Older Scots. In C. Jones (Ed.), *The Edinburgh History of the Scots Language* (pp. 112-155). Edinburgh: Edinburgh University Press.

- Montgomery, M. (1994). The evolution of verb concord in Scots. In A. Fenton, & D. A. MacDonald (Eds.), *Studies in Scots and Gaelic: Proceedings of the Third International Conference on the Languages of Scotland* (pp. 81-95). Edinburgh: Canongate Academic.
- Morris, R. (Ed.). (1874-1893). *Cursor Mundi: The Cursor o the World: A Northumbrian Poem of the XIVth Century in Four Versions*. London: Kegan Paul, Trench, Trübner & Co. for The Early English Text Society.
- Mossé, F. (1952). *A Handbook of Middle English*. (J. A. Walker, Trans.). Baltimore: The John Hopkins Press.
- Mourelatos, A. P. (1978). Events, processes, and states. *Linguistics and Philosophy*, 2, 415-434.
- Murray, J. A. (1873). *The Dialect of the Southern Counties of Scotland: Its Pronunciation, Grammar, and Historical Relations*. London: The Philological Society: Asher & Co.
- Mustanoja, T. F. (1960). *A Middle English Syntax: Part I: Parts of Speech*. Helsinki: Société Néophilologique.
- Mustanoja, T. F. (1983). Chaucer's use of *gan*: Some recent studies. In D. Gray, & E. G. Stanley (Eds.), *Middle English Studies: Presented to Norman Davis in Honour of his Seventieth Birthday* (pp. 59-64). Oxford: Clarendon Press.
- Narrog, H. (2005a). Modality, mood, and change of modal meanings: A new perspective. *Cognitive Linguistics*, 16 (4), 677-731.
- Narrog, H. (2005b). On defining modality again. *Language Sciences*, 27 (2), 165-192.
- Newmeyer, F. J. (1998) *Language Form and Language Function*. Cambridge, Massachusetts: The MIT Press.
- Norde, M. (2009). *Degrammaticalization*. Oxford: Oxford University Press.
- Nurmi, A. (1999). Auxiliary do in fifteenth-century English: Dialectal variation and formulaic use. In I. Taavitsainen, G. Melchers, & P. Pahta (Eds.), *Writing in Nonstandard English. Pragmatics & Beyond New Series*, 67 (pp. 225-242). Amsterdam: John Benjamins.

- Oakden, J. P. (1930). *Alliterative Poetry in Middle English: The Dialectal and Metrical Survey*. Publications of the University of Manchester No. CCV. English Series No. XVIII. Manchester: Manchester University Press.
- Oakes, M. P. (1998). *Statistics for Corpus Linguistics*. Edinburgh Textbooks in Empirical Linguistics. Edinburgh: Edinburgh University Press.
- Ogura, M. (1997). On the beginning and development of the *begin to* construction. In J. Fisiak (Ed.), *Studies in Middle English Linguistics*. Trends in Linguistics: Studies and Monographs 103 (pp. 403-428). Berlin: Mouton de Gruyter.
- Ogura, M. (1998). The grammaticalization in medieval English. In J. Fisiak, & M. Krygier (Eds.), *Advances in English Historical Linguistics* (pp. 293-314). Berlin: Mouton de Gruyter.
- Ogura, M. (2013). *Verbs in Medieval English: Differences in Verb Choice in Verse and Prose*. Berlin: Mouton de Gruyter.
- Oizumi, A. (1991-92). *A Complete Concordance to the Works of Geoffrey Chaucer*. Hildesheim-New York: Olms-Weidmann.
- Olsen, M. B. (1994). *A Semantic and Pragmatic Model of Lexical and Grammatical Aspect* (Doctoral dissertation, Northwestern University, Evanston, Illinois). Available from ProQuest Dissertations & Theses Global
<http://search.proquest.com.ezproxy.is.ed.ac.uk/pqdtglobal/docview/30414730/0/A5C088067D9F45F2PQ/5?accountid=10673>
- Omre, N. (1995). The culture of children in medieval England. *Past & Present*, 148, 48-88.
- Oxford English Dictionary Online*. (2014). (J. Pearsall, Editor, & Oxford University Press). Available from www.oed.com
- Pagliuca, W. (Ed.). (1994). *Perspectives on Grammaticalization*. Current Issues in Linguistic Theory, 109. Amsterdam: John Benjamins.
- Palmer, F. R. (1979). *Modality and the English Modals*. London: Longman.
- Palmer, F. R. (1986). *Mood and Modality*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- Palmer, F. R. ([1986] 2001). *Mood and Modality*. Cambridge Textbooks in Linguistics (2nd ed). Cambridge: Cambridge University Press.

- Panther, K.-U., & Radden, G. (Eds.). (1999). *Metonymy in Language and Thought. Human Cognitive Processing*, 4. Amsterdam: John Benjamins.
- Paul, H. ([1880] 2009). *Principien der Sprachgeschichte*. Cambridge Library Collection – Linguistics. Cambridge: Cambridge University Press.
- Pearsall, D. (1977). *Old English and Middle English Poetry*. London : Routledge & Kegan Paul.
- Pearsall, D. (1988). The development of Middle English romance. In D. Brewer (Ed.), *Studies in Medieval English Romances: Some New Approaches* (pp. 11-36). Cambridge: D. S. Brewer.
- Pearsall, D. (2011). The pleasure of popular romance: A prefatory essay. In R. Purdie, & M. Cichon (Eds.), *Medieval Romance, Medieval Contexts* (pp. 9-18). Cambridge: D. S. Brewer.
- Perry, G. G. (1867). *Religious Pieces in Prose and Verse: Edited from Robert Thornton's MS. (Cir. 1440) in the Lincoln Cathedral Library*. London: Early English Text Society.
- Piñón, C. (1997). Achievements in an event semantics. *Proceedings of Semantics and Linguistic Theory*, 7, 276-293.
- Plank, F. (1984). The modals story retold. *Studies in Language*, 8, 305-364.
- Poussa, P. (1990). A contact-universals origin for periphrastic *do*, with special consideration of OE-Celtic contact. In S. Adamson, V. A. Law, N. Vincent, & S. Wright (Eds.), *Papers from the 5th International Conference on English Historical Linguistics. Cambridge 6–9 April 1987*. Current Issues in Linguistic Theory, 65 (pp. 407-34). Amsterdam: John Benjamins.
- Preusler, W. (1956). Keltischer Einfluss im Englischen. *Indogermanische Forschungen*, 56, 178–191.
- Purdie, R. (2015). Medieval romance and the generic frictions of Barbour's Bruce. In S. Boardman, & S. Foran (Eds.), *Barbour's Bruce and its Cultural Contexts: Politics, Chivalry and Literature in Late Medieval Scotland* (pp. 51-74). Cambridge: D. S. Brewer.
- Putter, A. (2009). The metres and stanza forms of popular ballads. In R. L. Radulescu, & C. J. Rushton (Eds.), *A Companion to Medieval Popular Romance* (pp. 111-31). Cambridge: D. S. Brewer.

- Quirk, R., & Svartvik, J. (1970). Types and uses of non-finite clause in Chaucer. *English Studies*, 51, 393-411.
- Rayson, P. (1993-2014). *University of Lancaster: University Centre for Computer Corpus Research on Language: Log-likelihood and Effect Size Calculator Microsoft Excel Spreadsheet [Software]*. Retrieved from <http://ucrel.lancs.ac.uk/people/paul/SigEff.xlsx>
- Records of the Parliament of Scotland to 1707*. (2007-2017). University of St. Andrews. Retrieved from <http://www.rps.ac.uk/>
- Reeves, W. P. (1893). *A Study in the Language of Scottish Prose before 1600*. Baltimore: John Murphy.
- Richardson, P. K. (1991a). *Narrative Strategies in Early English Poetry* (Doctoral dissertation, University of California, Berkeley). Available from ProQuest Dissertations & Theses Global <https://search-proquestcom.ezproxy.is.ed.ac.uk/pqdtglobal/docview/303918023/83BBF20E3D40494FPQ/1?accountid=10673>
- Richardson, P. K. (1991b). Tense, discourse, and style. *Neuphilologische Mitteilungen*, XCII, 343-349.
- Riehle, W. (1981). *The Middle English Mystics*. (B. Standring, Trans.). London: Routledge and Kegan Paul.
- Rissanen, M. (1991). Spoken language and the history of *do*-periphrasis. In D. Kastovsky (Ed.), *Historical English Syntax*. Topics in English Linguistics, 2 (pp. 321-342). Berlin: Mouton de Gruyter.
- Rizzi, L. (1992). Reanalysis rules. In W. Bright (Ed.), *International Encyclopedia of Linguistics* (Vol. 3: 318-320). Oxford: Oxford University Press.
- Roberts, I., & Roussou, A. (2003). *Syntactic Change: A Minimalist Approach to Grammaticalization*. Cambridge: Cambridge University Press.
- Romaine, S. (1981). Contributions from Middle Scots syntax to a theory of syntactic change. In R. J. Lyall, & F. Riddy (Ed.), *Proceedings of the Third International Conference on Scottish Language and Literature: Medieval and Renaissance* (pp. 70-84). Stirling and Glasgow: University of Glasgow.

- Rothstein, S. (2008). *Explorations in Semantics: Structuring Events*. Oxford: Blackwell.
- Rubba, J. (1994). Grammaticalization as a semantic change. In W. Pagliuca (Ed.), *Perspectives on Grammaticalization*. Current Issues in Linguistic Theory, 109 (pp. 81-101). Amsterdam: John Benjamins.
- Sampson, G., & Churchill, R. C. ([1970] 1999). *The Concise Cambridge History of English Literature: Third edition / Revised throughout and with Additional Chapters on the Literature of the United States of America and the Mid Twentieth-Century Literature of the English-speaking World* by R. C. Churchill. Cambridge: Cambridge University Press.
- Sasse, H. J. (2006). Aspect and Aktionsart. In K. Allan (Ed.), *Concise Encyclopedia of Semantics* (pp. 31-34). Oxford: Elsevier.
- Saurer, W. (1984). *A Formal Semantics of Tense, Aspect and Aktionsarten*. Bloomington: Indiana University Linguistics Club.
- Scahill, J., & Rogerson, M. (2005). *Annotated Bibliographies of Old and Middle English Literature: Middle English Saints' Legends* (Vol. 3). Cambridge: D. S. Brewer.
- Schendl, H. (2000). The third person present plural in Shakespeare's First Folio: A case of interaction of morphology and syntax. In C. Dalton-Puffer, & N. Ritt (Eds.), *Trends in Linguistics. Words: Structure, Meaning, Function* (pp. 263-276). Berlin: Mouton de Gruyter.
- Scott, M. (2004). Use of Scottish place-names as evidence in historical dictionaries. In C. Kay, C. Hough, & I. Wotherspoon (Eds.) *New Perspectives on English Historical Linguistics. Selected Papers from 12 ICEHL, Glasgow, 21-26 August 2002*. Current Issues in Linguistic Theory, 252 (pp. 213-224). Amsterdam: John Benjamins.
- Silva-Corvalán, C. (1983). Tense and aspect in oral Spanish narrative: Context and meaning. *Language*, 59 (4), 760-780.
- Simpson, P. (2004). *Stylistics: A Resource Book for Students*. London: Routledge.
- Sims, L. (2008). *The Grammaticalization of Ingressive Aspect in Early English* (Doctoral dissertation, Arizona State University). Available from ProQuest Dissertations & Theses Global

<http://search.proquest.com.ezproxy.is.ed.ac.uk/pqdtglobal/docview/30468620/7/51C0677FDEDA452FPQ/1?accountid=10673>

- Sims, L. (2014). Aspectual loss and renewal: *onginnan*, *beginnan* and *start*. In I. Hegedűs, & D. Pődör (Eds.), *Periphrasis, Replacement and Renewal: Studies in English Historical Linguistics* (pp. 57-82). Newcastle upon Tyne: Cambridge Scholars.
- Sinclair, J. (1995). Tipologia textual EAGLES. *Actas Do XI Encontro Nacional Da Associação Portuguesa de Linguística, 1*, 39-92. Retrieved from http://www.apl.org.pt/docs/actas-11-encontro-apl-1995_voll1.pdf
- Sinclair, J. (1996). *EAGLES: Preliminary recommendations on corpus typology*. EAG-- TCWG--CTYP/P. Retrieved from <http://www.ilc.cnr.it/EAGLES/corpus/typ/corpus.html>
- Skeat, W. W. (1915). *The Complete Works of Geoffrey Chaucer*. Oxford: Clarendon Press.
- Smith, C. S. (1991). *The Parameter of Aspect*. London: Kluwer Academic Publishers.
- Smith, C. S. (1997). *The Parameter of Aspect* (2nd ed.). London: Kluwer Academic Publishers.
- Smith, J. (1996). *An Historical Study of English: Function, Form and Change*. London: Routledge.
- Smith, J. (2000). Scots. In G. Price (Ed.), *Languages in Britain and Ireland* (pp. 159-170). Oxford: Blackwell.
- Smith, J. (2013). Textual afterlives: Barbour's Bruce and Hary's Wallace. In J. M. Kirk, & I. Macleod (Eds.), *Scots: Studies in its Literature and Language* (pp. 37-70). Amsterdam: Rodopi B. V.
- Smitterberg, E. (2005). *The Progressive in 19th-Century English: A Process of Integration*. Amsterdam: Rodopi.
- Smyser, H. M. (1967). Chaucer's use of *gin* and *do*. *Speculum*, 42 (1), 68-83.
- Spinks, L. (2010). Metre and rhythm. In D. Cavanagh, A. Gillis, M. Keown, J. Loxley, & R. Stevenson (Eds.), *The Edinburgh Introduction to Studying English Literature* (pp. 47-56). Edinburgh: Edinburgh University Press.
- Steele, S. (1975). Is it possible? *Working Papers on Language Universals*, 18, 35-58.

- Stern, G. (1931). *Meaning and Change of Meaning: With Special Reference to the English Language*. Göteborg: Elanders Boktryckeri Aktiebolag.
- Stevens, M. (1981). Did the Wakefield master write a nine-line stanza? *Comparative Drama*, 15 (2), 99-119.
- Stevens, M., & Cawley, A. C. (Eds.). (1994). *The Towneley Plays*. Oxford: Oxford University Press.
- Strachan, J., & Terry, R. ([2000] 2011). *Poetry* (2nd ed.). Edinburgh: Edinburgh University Press.
- Sweet, H. ([1898] 2014). *A New English Grammar: Logical and Historical. Volume 2: Syntax*. Cambridge Library Collection – Linguistics. Cambridge: Cambridge University Press.
- Sweetser, E. E. (1988). Grammaticalization and semantic bleaching. *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society* (pp. 389-405). The Annual Proceedings of the Berkeley Linguistics Society.
- Sweetser, E. E. (1990). *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge Studies in Linguistics, 54. Cambridge: Cambridge University Press.
- Tajima, M. (1975). The Gawain-poet's use of 'con'. *Neuphilologische Mitteilungen*, LXXVI, 429-438.
- Taylor, R. (1917). Some notes on the use of *can* and *couth* as preteritive auxiliaries in Early and Middle Scottish poetry. *The Journal of English and Germanic Philology*, 16 (4), 573-591.
- Teaching Association for Medieval Studies (Middle English Text Series Texts Online)*. (n.d.). (R. A. Peck, Editor, T. R. Library, Compilers). The University of Rochester. Retrieved from <http://d.lib.rochester.edu/teams>
- Tenny, C. (1994). *Aspectual Roles and the Syntax-Semantics Interface*. London: Kluwer Academic.
- Terasawa, Y. (1974). Some notes on ME *gan* periphrasis. *Poetica*, 89-105. *The Breadalbane Collection*. (n.d.). The University of Edinburgh. Available from <http://www.ed.ac.uk/divinity/research/resources/breadalbane>

- The Corpus of Early English Correspondence*. (1993-). (T. Nevalainen, Project Leader) VARIENG: The Research Unit for Variation, Contacts and Change in English. University of Helsinki.
- The Corpus of English Dialogues*. (n.d.). (M. Kytö, & J. Culpeper, Compilers, in collaboration with T. Walker, & D. Archer) Uppsala University and Lancaster University.
- The Corpus of Scottish Correspondence*. (2007). (A. Meurman-Solin, Compiler) Department of Modern Languages, University of Helsinki.
- The Helsinki Corpus of English Texts*. (1991). (M. Rissanen, M. Kytö, L. Kahlas-Tarkka, M. Kilpiö, S. Nevanlinna, & I. Taavitsainen, et al., Compilers) Helsinki: Department of Modern Languages, University of Helsinki.
- The Helsinki Corpus of Older Scots*. (1995). (A. Meurman-Solin, Compiler) Helsinki: Department of Modern Languages, University of Helsinki.
- The Holy Bible: King James Version*. (2011). Peabody, MA: Hendrickson Marketing, LLC.
- The Innsbruck Corpus of Middle English Prose (Sampler)*. (1992-97). Innsbruck Computer Archive of Machine-Readable English Texts. (M. Markus, Compiler) Innsbruck: Universität Innsbruck.
- The Lampeter Corpus of Early Modern English Tracts*. (1991-). (C. Claridge, & R. Siemund, Compilers). English Department: Technische Universität Chemnitz. Available from <https://www.tuchemnitz.de/phil/english/sections/linguist/real/independent/lampeter/tei/Lampeter.html>
- The Penn-Helsinki Parsed Corpus of Middle English (PPCME2)*. (2000). (A. Kroch, & A. Taylor, Compilers). Department of Linguistics, University of Pennsylvania. Available from <http://www.ling.upenn.edu/ppche-release-2016/PPCME2-RELEASE-4>
- Tieken-Boon van Ostade, I. (1985). *Do*-support in the writings of Lady Mary Wortley Montagu: A change in progress. *Folia Linguistica Historica*, 6, 127-151.
- Tieken-Boon van Ostade, I. (1987). *The Auxiliary 'Do' in Eighteenth-Century English: A Sociohistorical-Linguistic Approach*. Dordrecht: JCG Printing.

- Tieken-Boon van Ostade, I. (1989). Dr. Johnson and the auxiliary *do*. *Folia Linguistica Historica*, 10, 145-62.
- Tieken-Boon van Ostade, I. (1990). The origin and development of periphrastic auxiliary *do*: A case of destigmatization. *Nowele*, 16, 3-52.
- Timberlake, A. (1977). Reanalysis and actualization in syntactic change. In C. N. Li (Ed.), *Mechanisms of Syntactic Change* (pp. 141-180). Austin: University of Texas Press.
- Tobin, Y. (1993). *Aspect in English Verb: Process and Result in Language*. London: Longman.
- Tonkin, E. (1992). *Narrating our Pasts: The Social Construction of Oral History*. Cambridge Studies in Oral and Literate Culture. Cambridge: Cambridge University Press.
- Traugott, E. C. (1972). *A History of English Syntax: A Transformational Approach to the History of English Sentence Structure*. New York: Holt, Rinehart, & Winston, Inc.
- Traugott, E. C. (1995). The role of the development of discourse markers in a theory of grammaticalization: Paper presented at ICHL XII, Manchester 1995. Retrieved from Elizabeth C. Traugott: Papers Available On-line <http://www.stanford.edu/~traugott/papers/discourse.pdf>
- Traugott, E. C. (2010). Revisiting subjectification and intersubjectification. Retrieved from Elizabeth C. Traugott: Papers Available On-line <http://www.stanford.edu/~traugott/resources/TraugottDavidseIntersbfn.pdf>
- Traugott, E. C. (2012a). On the persistence of ambiguous linguistic contexts over time: Implications for corpus research on micro-changes. In J. Mukherjee, & M. Huber (Eds.), *Corpus Linguistics and Variation in English: Theory and Description* (pp. 231-246). Amsterdam: Rodopi.
- Traugott, E. C. (2012b). The status of onset contexts in analysis of micro-changes. In M. Kyto (Ed.), *Language and Computers, English Corpus Linguistics: Crossing Paths* (pp. 221-255). Amsterdam: Rodopi.
- Traugott, E. C., & Heine, B. (1991). Introduction. In E. C. Traugott, & B. Heine (Eds.), *Approaches to Grammaticalization*. Typological Studies in Language 19: 1 (pp. 1-14). Amsterdam: John Benjamins.

- Traugott, E. C., & König, E. (1991). The semantics-pragmatics of grammaticalization revisited. In E. C. Traugott, & B. Heine (Eds.), *Approaches to Grammaticalization*. Typological Studies in Language 19: 1 (pp. 189-218). Amsterdam: John Benjamins.
- Traugott, E. C., & Trousdale, G. (Eds.). (2010). *Gradience, Gradualness and Grammaticalization*. Typological Studies in Language, 90. Amsterdam: John Benjamins.
- Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and Constructional Changes*. Oxford Studies in Diachronic and Historical Linguistics. Oxford: Oxford University Press.
- Turco, L. (2000). *The Book of Forms: A Handbook of Poetics* (3rd ed.). London: University Press of New England.
- Turville-Petre, T. (1977). *The Alliterative Revival*. Cambridge: D. S. Brewer.
- Ullmann, S. (1963). *The Principles of Semantics* (2nd ed.). Oxford: Blackwell.
- Van der Auwera, J. (2002). More thoughts on degrammaticalization. In I. Wischer, & G. Diewald (Eds.), *New Reflections on Grammaticalization*. Typological Studies in Language, 49 (pp. 19-29). Amsterdam: John Benjamins.
- Van Gelderen, E. (2004). *Grammaticalization as Economy*. *Linguistik Aktuell/Linguistics Today*, 71. Amsterdam: John Benjamins.
- Van Kemenade, A., & Los, B. (2003). Particles and prefixes in Dutch and English. In G. Booij, & J. van Marle (Eds.), *Yearbook of Morphology 2003* (pp. 79-117). Dordrecht: Springer.
- Van Peer, W. (1986). *Stylistics and Psychology: Investigations of Foregrounding*. London: Croom Helm.
- Van Peer, W., Zyngier, S., & Hakemulder, J. (2007). Foregrounding: Past, present, future. In D. L. Hoover, & S. Lattig (Eds.), *Stylistics: Prospect & Retrospect* (pp. 1-22). Amsterdam: Rodopi.
- Vendler, Z. (1957). Verbs and times. *The Philosophical Review*, 143-160.
- Vendler, Z. (1967). *Linguistics in Philosophy*. Ithaca, New York: Cornell University Press.
- Verkuyl, H. J. (1972). *On the Compositional Nature of the Aspects*. Dordrecht: Reidel.

- Viana, V., Zyngier, S., & Barnbrook, G. (Eds.). (2011). *Perspectives on Corpus Linguistics*. Studies in Corpus Linguistics, 48. Amsterdam: John Benjamins.
- Visser, F. T. (1964). The 'historical present' in Middle English verse narratives. In *English Studies: Presented to R.W. Zandvoort on Occasion of his Seventieth Birthday*. English Studies, 45 (pp. 135-142). Amsterdam: Swets and Zeitlinger.
- Visser, F. T. (1969). *An Historical Syntax of the English Language: Part III: First Half: Syntactical Units with Two Verbs*. Leiden: E. J. Brill.
- Visser, F. T. (1972). *An Historical Syntax of the English Language: Part II: Syntactical Units with One Verb*. Leiden: E. J. Brill.
- Visser, F. T. (1973). *An English Syntax of the English Language: Part III: Second Half: Syntactical Units with Two and with More Verbs*. Leiden: E. J. Brill.
- Von Contzen, E. (2016). *The Scottish Legendary: Towards a Poetics of Hagiographic Narration*. Manchester: Manchester University Press.
- Ward, A. W., Waller, A. R., Trent, W. P., Erskine, J., Sherman, S. S., & van Doren, C. (Eds.). ([1907-21] 2000). *The Cambridge History of English and American Literature in 18 Volumes* (Bartleby.com ed., Vol. 2). Cambridge: Cambridge University Press.
- Warner, A. R. (1993). *English Auxiliaries: Structure and History*. Cambridge: Cambridge University Press.
- Weber, H. (Ed.). (1810). *Metrical Romances of the Thirteenth, Fourteenth, and Fifteenth Centuries: Published from Ancient Manuscripts. With an Introduction, Notes, and a Glossary* (Vol. 1). Edinburgh: Archibald Constable and Co.
- Wegener, H. (2002). The evolution of the German modal particle *denn*. In I. Wischer, & G. Diewald (Eds.), *New Reflections on Grammaticalization*. Typological Studies in Language, 49 (pp. 379-394). Amsterdam: John Benjamins.
- Whatley, G. E., Thompson, A. B., & Upchurch, R. K. (Eds.). (2004a). *St. Andrew and the Three Questions: Introduction. Saints' Lives in Middle English Collections* (University of Rochester). Retrieved from TEAMS: Middle English Texts Series <http://d.lib.rochester.edu/teams/text/whatley-saints->

[lives-in-middle-english-collections-st-andrew-and-the-three-questions-introduction](#)

- Whatley, G. E., Thompson, A. B., & Upchurch, R. K. (Eds.). (2004b). *The Life of St. Julian the Hospitaller: Introduction. Saints' Lives in Middle English Collections* (University of Rochester). Retrieved from TEAMS: Middle English Texts Series <http://d.lib.rochester.edu/teams/text/whatley-saints-lives-in-middle-english-collections-life-of-st-julian-hospitaller-introduction>
- Whetter, K. S. (2008). *Understanding Genre and Medieval Romance*. Farnham: Ashgate.
- Whitehead, C. (2005). Middle English religious lyrics. In T. G. Duncan (Ed.), *A Companion to the Middle English Lyric* (pp. 96-119). Cambridge: D. S. Brewer.
- Williamson, K. (2002). The dialectology of 'English' north of the Humber, c. 1380-1500. In T. Fanego, B. Méndez-Naya, & E. Seoane (Eds.), *Sounds, Words, Texts and Change: Selected Papers from 11 ICEHL, Santiago de Compostela, 7-11 Septemeber 2000*. *Current Issues in Linguistic Theory*, 224 (pp. 253-286). Amsterdam: John Benjamins.
- Williamson, K. (2008). *The LAOS Manual* (University of Edinburgh). Retrieved from http://www.lel.ed.ac.uk/ihd/laos1/laos1_manual.html
- Wittig, S. (1978). *Stylistic and Narrative Structures in the Middle English Romances*. Austin: University of Texas Press.
- Wright, J. (1898). *The English Dialect Dictionary being the Complete Vocabulary of All Dialect Words Still in Use, or Known to Have Been in Use During the Last Two Hundred Years: Founded on the Publications of the English Dialect Society and on a Large Amount of Material Never Before Printed* (Vol. 1: A-C). London: Henry Frowde, Amen Corner.
- Wright, L. (2004). Life after degrammaticalization: Plural *be*. In H. Lindquist, & C. Mair (Eds.), *Corpus Approaches to Grammaticalization in English*. *Studies in Corpus Linguistics*, 13 (pp. 211-226). Amsterdam: John Benjamins.
- Wuth, A. (1915). *Aktionsarten der Verba bei Cynewulf*. Weida i. Thür: Thomas & Hubert.

- Wyld, H. C. (1914). *A Short History of English with a Bibliography of Recent Books on the Subject, and Lists of Texts and Editions*. London: John Murray.
- Zilling, O. (1918). *Das Hilfsverb DO im Mittelenglischen*. Inaugural-Dissertation. Halle (Saale): Ehrhardt Karras.
- Zubizarreta, M. L. (1985). The relation between morphophonology and morphosyntax: The case of Romance causatives. *Linguistic Inquiry*, 16 (2), 247-289.