UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS

IZABELLA ROSA MALTA

THE USES AND MEANINGS OF THE MIDDLE ENGLISH VERB *HAVEN*:
A CORPUS-BASED STUDY

BELO HORIZONTE
2020

IZABELLA ROSA MALTA

THE USES AND MEANINGS OF THE MIDDLE ENGLISH VERB *HAVEN*:

A CORPUS-BASED STUDY

Dissertação apresentada ao Programa de Pós-Graduação em Estudos Linguísticos da Faculdade de Letras da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do título de Mestre em Linguística Teórica e Descritiva.

Área de Concentração: Linguística Teórica e Descritiva
Linha de Pesquisa: (1D) Estudos Linguísticos baseados em Corpora
Orientadora: Profª. Drª. Heliana Ribeiro de Mello

Belo Horizonte
Faculdade de Letras da UFMG
2020

Ficha catalográfica elaborada pelos Bibliotecários da Biblioteca FALE/UFMG

Malta, Izabella Rosa.

M261u The uses and meanings of the middle english verb *haven* [manuscrito] : a corpus-based study / Izabella Rosa Malta. – 2020.

199 f., enc. : il., grafs (color), tabs (color)

Orientadora: Heliana Ribeiro de Mello.

Área de concentração: Linguística Teórica e Descritiva.

Linha de Pesquisa: Estudos Linguísticos Baseados Em Corpora.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Faculdade de Letras.

Bibliografia: f. 131-141.

Apêndices: f. 142-199.

1. Linguística histórica – Teses. 2. Linguística de corpus – Teses. 3. Lingua inglesa – Verbos – Teses. I. Mello, Heliana. II. Universidade Federal de Minas Gerais. Faculdade de Letras. III. Título.

CDD: 410

FOLHA DE APROVAÇÃO

The uses and meanings of the Middle English verb "haven": a corpus-based study

IZABELLA ROSA MALTA

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ESTUDOS LINGUÍSTICOS, como requisito para obtenção do grau de Mestre em ESTUDOS LINGUÍSTICOS, área de concentração LINGUÍSTICA TEÓRICA E DESCRITIVA, linha de pesquisa Estudos Linguísticos Baseados em Corpora.

Aprovada em 17 de fevereiro de 2020, pela banca constituída pelos membros:

Prof(a). Heliana Ribeiro de Mello - Orientadora
UFMG

Prof(a). Giulia Bossaglia
UFMG

Prof(a). Aparecida de Araújo Oliveira
UFV

Belo Horizonte, 17 de fevereiro de 2020.

Profa. Ana Larissa Adorno Marciotto Oliveira
Subcoord^a. Programa de Pós- Graduação
em Estudos Linguísticos
FALE/UFMG

(*this page intentionally left blank*)

ACKNOWLEDGMENTS

To Eniara and Marcelino, my parents, who provided me everything and cared for my education. Without them and their support, I would not have come so far.

To Rodrigo Araújo e Castro (*aka* Rodrigato), my cat-like Datawizard and husband-to-be. Whenever I sought counsel he was there, honouring me with his presence and bright knowledge.

To Thor, the sweetest cat in this world, named after the God of Thunder. This little creature is my everyday joy.

To my friends. A special thank-you to Adriana Alves, Katherine Oliva and Raquel Rossini, who, alongside Rodrigato, supported me since the beginning of this quest.

Last but not least, to Professor Heliana Mello, the wisest person and advisor I have ever known.

ABSTRACT

Historical Linguistics research approaches have changed drastically from document-based studies to corpus-based studies with modern methods to investigate diachronic linguistic phenomena, such as the linguistic change of Middle English verbs. This study aims to investigate the syntactic-semantic behaviour of the Middle English verb *haven* ('to have'). This study draws on the interface between Corpus Linguistics (henceforth CL) and Historical Linguistics (henceforth HL) through a syntactic-semantic viewpoint on verb valency (Perini, 2015). For this purpose, CMEL – *Corpus of Middle English Lyrics* was compiled, comprehending lyrics from the timespan between the 12th to the 15th centuries. The corpus analysis was conducted according to the following procedures: a) selecting philological editions of lyrics; b) creating one subcorpus for each century; c) processing data with *UAM CorpusTool* 3.3m and *R* (R CORE, 2019); d) filtering the lexical occurrences of *haven*, and e) sampling them for each century. Based on a CMEL's sample of lexical occurrences, an in-depth discussion from a diachronic point of view suggested that the syntactic-semantic behaviour of *haven*, more specifically its set of valency and argument structures, has changed over the covered centuries. This resulted in more syntactic fixity and more specific uses and meanings of *haven*. In other words, *haven* had two uses, grammatical and lexical, the latter relevant to show that different meanings arose. In conclusion, linguistic change in respect to meanings and verb valency of *haven* was a matter of fact in this research. This work will benefit not only interfaces of CL with HL, but also other studies investigating different word classes and phenomena.

Keywords: Historical Linguistics; Corpus Linguistics; Middle English; verb valency.

RESUMO

As abordagens dos estudos no campo da Linguística Histórica mudaram drasticamente, não estando mais restritas aos documentos históricos devido a métodos contemporâneos baseados em corpora para a investigação de fenômenos linguísticos diacrônicos como a mudança linguística de verbos do Inglês Médio. O presente estudo visa investigar o comportamento sintático-semântico do verbo em Inglês Médio *haven* ('ter'). Este estudo se ancora na interface entre a Linguística de Corpus (LC) e a Linguística Histórica (LH) por meio de uma perspectiva sintático-semântica em valência verbal (PERINI, 2015). Para tal, o CMEL – *Corpus of Middle English Lyrics* ("Corpus de Líricas em Inglês Médio") foi compilado, compreendendo poesias líricas datadas entre os séculos XII e XV. A análise do corpus foi conduzida a partir dos seguintes passos: a) selecionar edições filológicas das líricas; b) criar um subcorpus referente a cada século; c) processar os dados nos softwares *UAM CorpusTool* 3.3m e *R* (R CORE, 2019); d) filtrar as ocorrências lexicais de *haven*, e e) amostrar tais ocorrências para cada século. Baseada na amostra de ocorrências lexicais do CMEL, uma discussão aprofundada de um ponto de vista diacrônico indicou que o comportamento sintático-semântico de *haven*, mais especificamente seu conjunto de valências verbais e estruturas argumentais, mudou no decorrer dos séculos. Tal processo de mudança resultou em uma sintaxe mais fixa, assim como em usos e significados mais específicos. Em outras palavras, o verbo *haven* apresentou dois usos, gramatical e lexical, sendo este último relevante para mostrar o surgimento de significados distintos. Concluiu-se que a mudança linguística no que se refere aos significados e valências verbais de *haven* pôde ser atestada neste estudo. Este trabalho beneficiará não só a interface entre a LC e a LH, mas também outros estudos que investiguem outras classes de palavra e fenômenos linguísticos.

Palavras-chave: Linguística Histórica; Linguística de Corpus; Inglês Médio; valência verbal.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF GRAPHS

LIST OF TEXTS OF CMEL

12th CENTURY

A001 - Proverbs of Alfred i - Incipiunt documenta Regis Aluredi
A002 - Proverbs of Alfred ii
A003 - Proverbs of Alfred iii
A004 - Proverbs of Alfred iv
A005 - Proverbs of Alfred v
A006 - Proverbs of Alfred vi
A007 - Proverbs of Alfred vii
A008 - Proverbs of Alfred viii
A009 - Proverbs of Alfred ix
A010 - Proverbs of Alfred x
A011 - Proverbs of Alfred xi
A012 - Proverbs of Alfred xii
A013 - Proverbs of Alfred xiii
A014 - Proverbs of Alfred xiv
A015 - Proverbs of Alfred xv
A016 - Proverbs of Alfred xvi
A017 - Proverbs of Alfred xvii
A018 - Proverbs of Alfred xviii
A019 - Proverbs of Alfred xix
A020 - Proverbs of Alfred xx
A021 - Proverbs of Alfred xxi
A022 - Proverbs of Alfred xxii
A023 - Proverbs of Alfred xxiii - Expliciunt dicta Regis Aluredi
A024 - Worcester Fragments A - Sicut Oues absque Pastore
A025 - Worcester Fragments B - Homo Putredo et Filius Hominis Vermis
A026 - Worcester Fragments C - Confabulator cum Amaritudine Animae
A027 - Poema Morale
A028 - Ormulum - Preface
A029 - Ormulum - Introduction
A030 - Ormulum - Secundum Lucam I
A031 - Durham
A032 - Ormulum - Secundum Lucam XV

13th CENTURY

B001 - The milde Lamb, y-sprad o rode
B002 - Worldes blis ne last throwe
B003 - Swarte smeked smithes
B004 - Love Song
B005 - Blow, Northern Wind
B006 - Edi be thou, hevenë queen

B007 - Havelok the Dane
B008 - Judas
B009 - King Horn
B010 - Lavedy seynte Marie
B011 - Levëdie, Ich thonkë thee
B012 - Now this foulës singëth
B013 - Of on that is so fayr and bright
B014 - On hire is al mi lif y-long
B015 - Somer is come and winter gon
B016 - When the Nightingale Sings
B017 - Love is soft, love is swet
B018 - Though I can wittës ful i-wis
B019 - Foulës in the frith
B020 - Where ben they before us were
B021 - I sike al when I singe
B022 - Hit was upon a Shere Thorsday
B023 - Man mai longe him livës wenë
B024 - Whan I thenkë thingës thre
B025 - If man him bithoughtë
B026 - Whan mine eyen mistëth
B027 - Whan the turuf is thy tour
B028 - Dorë, go thou stillë
B029 - Ich have y-don al myn youth
B030 - The Orison of our Lady
B031 - Layamon
B032 - No more ne wil I wiked be
B033 - Swetë Jhesu, king of blisse
B034 - Gabriel, from hevenë king
B035 - The Bestiary - Natura leonis i^{a}
B036 - The Bestiary - Natura leonis ij^{a}
B037 - The Bestiary - Natura leonis iij^{a}
B038 - The Bestiary - Significacio prime nature
B039 - The Bestiary - Significacio prime nature ij^{a} - iij^{a}
B040 - The Bestiary - Natura aquile
B041 - The Bestiary - Natura aquile - Significacio
B042 - The Bestiary - Natura serpentis j^{a}
B043 - The Bestiary - Natura serpentis ij^{a}
B044 - The Bestiary - Natura serpentis - Significacio
B045 - The Bestiary - Natura formice
B046 - The Bestiary - Natura formice - Significacio
B047 - The Bestiary - Natura cerui
B048 - The Bestiary - Natura cerui - Significacio prima
B049 - The Bestiary - Natura cerui - Natura ij^{a}
B050 - The Bestiary - Natura cerui - Significacio ij^{a}
B051 - The Bestiary - Natura Wulpis
B052 - The Bestiary - Natura Wulpis – Significacio

B053 - The Bestiary - Natura Wulpis - Significacio 2
B054 - The Bestiary - Natura aranéé
B055 - The Bestiary - Natura aranéé - Significacio
B056 - The Bestiary - Natura cetegrandie
B057 - The Bestiary - Natura cetegrandie - Significacio
B058 - The Bestiary - Natura sirene
B059 - The Bestiary - Natura sirene - Significacio
B060 - The Bestiary - Natura elephantis
B061 - The Bestiary - Natura elephantis - Significacio
B062 - The Bestiary - Natura turturis
B063 - The Bestiary - Natura turturis - Significacio
B064 - The Bestiary - Natura pantere
B065 - The Bestiary - Natura pantere - Significacio
B066 - The Bestiary - Natura columbe - significacio
B067 - The Owl and the Nightingale
B068 - Now goth sunnë under wode
B069 - Whyt is thy naked brest
B070 - Jesu Cristës mildë moder
B071 - Stond wel, moder, under rode

14th CENTURY

C001 - Now welcome, somer, with thy sonnë softe
C002 - The Complaint of Chaucer to his Purse
C003 - Womanly Noblesse
C004 - What is he, this lordling
C005 - With longing I am lad
C006 - A waylë whyte as whallës bon
C007 - Ichot a byrde in a bour
C008 - Litel wot it any man
C009 - Ichot a byrde in boure bryght
C010 - Bitwenë March and Avëril
C011 - In May it mirieth when it dawës
C012 - In a fryth as I gan farë fremëde
C013 - My deth I love, my lyf Ich hatë
C014 - Moste I ryde by Rybbësdale
C015 - Weping hath myn wongës wet
C016 - Hye Loverd, thou here my bone
C017 - Nou shrinkëth rose and lylie-flour
C018 - Middelerd for man was mad
C019 - As I me rod this ender day
C020 - Ich herde men upon mold
C021 - Man in the moonë stont and strit
C022 - Wynnere and Wastoure - Introduction
C023 - Wynnere and Wastoure - Fitt 1
C024 - Wynnere and Wastoure - Fitt 2

C025 - Wynnere and Wastoure - Fitt 3
C026 - Worldes blisse, have god day!
C027 - Jhesus doth him bymene
C028 - Alliterative Morte Arthure - Arthur's Farewell to Guenevere
C029 - Were ther outher in this toun
C030 - Of every kinnë tre
C031 - Al night by the rosë, rosë
C032 - Al gold, Janet, is thin her
C033 - As I stod on a day
C034 - Lullay, lullay, litel child, why wepëstou so sore
C035 - Lenten ys come with love to toune
C036 - I woldë wite of sum wys wight
C037 - To Rosëmounde
C038 - Canticus Troili
C039 - Against women unconstant
C040 - Complaynt D'Amours
C041 - Ler to love as I love thee
C042 - Why have ye no routhe on my child
C043 - My folk, now answerë me
C044 - A sory beverage it is
C045 - D…dronken
C046 - Lullay, lullay, litel child [2]
C047 - Lullay, lullay, litel child [1]
C048 - Lovely ter of lovely eye
C049 - Ye that pasen be the wey
C050 - Skottes out of Berwik
C051 - Whan men ben meriest at her mele
C052 - The Pearl - The New Jerusalem
C053 - Sir Gawain and the Green Knight - Gawain's Journey
C054 - Sir Gawain and the Green Knight - 2 The Stag-hunt and Gawain's First Temptation

15th CENTURY

D001 - Dear son, leave thy weeping
D002 - Here I Sit Alone, Alas! Alone
D003 - It is my Father's Will
D004 - He said Ba-Bay, she said Lullay
D005 - She Sang, Dear son, Lullay
D006 - Filius Regis Mortuus Est
D007 - An Appeal to all Mothers
D008 - O Thou, with Heart of Stone
D009 - Who cannot Weep come Learn of me
D010 - Our Lady's Imprecation

D011 - Stabat Mater Dolorosa
D012 - Thou shalt Bear the Fruit of Life
D013 - A Salutation to the Virgin
D014 - Ave Gloriosa
D015 - Ave Regina Celorum, I
D016 - Ave Regina Celorum, II
D017 - Regina Celi Letare, II
D018 - Regina Celi Letare, III
D019 - The Five Joys of Our Lady
D020 - The Five Joys of Our Lady, with Acrostic
D021 - A salutation by the Heavenly Joys
D022 - A Song of the Assumption
D023 - A Prayer to the Three Persons on the Trinity
D024 - In One is All
D025 - God Governs for the Best
D026 - Christ, Defend Me from My Enemies
D027 - A Prayer by the Wounds against the Deadly Sins
D028 - Gabriel Came Down with Light
D029 - Verbum Caro Factum Est
D030 - Brother, Abide
D031 - Medicines to Cure the Deadly Sins
D032 - The Mourners at the Cross
D033 - Thou Sinful Man that by Me Goes
D034 - Why Art Thou, Man, Unkind
D035 - A Dialogue between Natura Hominis and Bonitas Dei
D036 - A Series of Triads
D037 - Speed Our King on His Journey
D038 - God Send Us Patience in Our Old Age
D039 - Balthazer, Melchior and Jasper

ACRONYMS

A - Adjunct
ADJ - Adjective
AdjP – Adjective Phrase
AdvP – Adverb Phrase
ALC - Allocutive
Arg1 – Argument 1
Arg2 – Argument 2
Arg3 – Argument 3
CEEC – *Corpus of Early English Correspondence*
CL - Clause
CLP – *Corpus Lexicográfico do Português*
CMEL – *Corpus of Middle English Lyrics*
CoL – Corpus Linguistics
EME – Early Modern English
Ger. – German (*Hochdeutsch*)
H.L. – Hapax legomena
HL – Historical Linguistics
IcePaHC – *Icelandic Parsed Historical Corpus*
IMP - Imperative
IND - Indicative
INF - Infinitive
ME – Middle English
MED – Middle English Dictionary
N – Noun
NEG – Negation
NP – Noun Phrase
O – Object
OE – Old English
PCMEP – *Parsed Corpus of Middle English Poetry*
PDE – Present-Day English
PL – Plural
PP – Prepositional Phrase
PPCEME – *Penn-Helsinki Parsed Corpus of Early Modern English*
PPCMBE2 – *Penn-Helsinki Parsed Corpus of Modern British English* (second edition)
PPCME2 – *Penn-Helsinki Corpus of Middle English* (second edition)
PRS – Present tense
PST – Past tense
PTCP – Participle
S – Subject
SBJV – Subjunctive
SG – Singular
TMN – Time, Mode and Negation
V - Verb
WALS – World Atlas of Linguistic Structures

INTRODUCTION

The aim of this study lies on an in-depth investigation of the syntactic-semantic behaviour of the Middle English verb *haven* ("to have") based on a corpus designed specifically for this purpose. Before introducing the topic alongside its main and specific objectives, some considerations on the history of the English language are more than essential to serve as a background. Being so, I will subdivide the Introduction into three Sections, namely Section 1, bringing forward how English evolved as a legitimate language in the Middle Ages; Section 2 will next explain the general and main objectives of this study, and, finally, Section 3 will present the structure of the study.

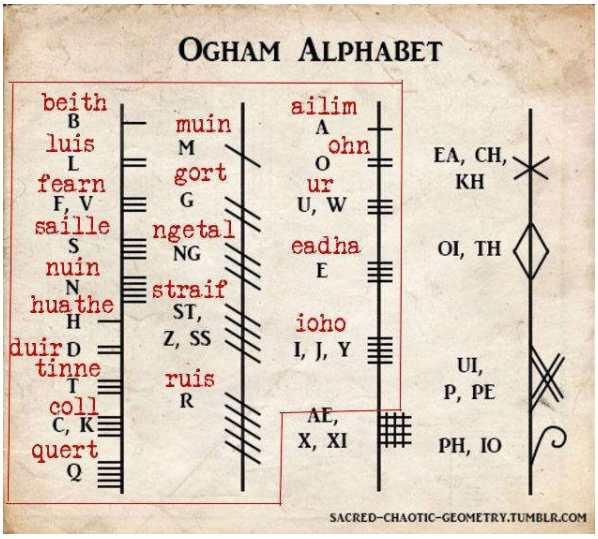
Section 1 – The history of English

Let us begin this Section with a core question: *what is the English language?* For a lay audience, it may be defined simply as the language of globalisation spoken all over the world, may it be the official language of influent nations such as the United States, or may it be learned in most known speech communities as a second or foreign language for a variety of purposes – business, education, tourism, among others. The supposed definition given by a lay audience is not wrong though; it must yet be deepened, taking into consideration that a linguist's point of view tells a very different story. For Linguistics students or lecturers, a well-known fact is that English is an Indo-European language, belonging to the Germanic *genus* according to WALS – *The World Atlas of Language Structures Online*[1]. As any other natural language, English has its own history.

[1] Available at: https://wals.info/

Here begins a bit of storytelling. Before this language became well-known in many fields as it is nowadays, it used to be quite unprivileged. In the lands currently known as Great Britain, Scotland, Ireland, Cornwall, Wales and the Isle of Man, a variety of languages belonging to the Celtic *genus* (also Indo-European) was spoken and registered with their own writing systems. The most famous is the alphabet called *ogham* to write Old Irish (c. 4th-10th centuries). The Celts, very widespread over Europe, "also spoke Latin, the language of the Roman Empire" (Hogg, 2002, p. 1), but *ogham* (see Figure 1) was perhaps unknown to the Romans. These lands had internal and external conflicts, not the focus of this introduction. The Celtic branch could be subdivided into languages, some of them still alive (e.g. Welsh, Irish and Scottish Gaelic in some communities in their respective countries) and documented and others which have unfortunately perished over the centuries.

Figure 1. Ogham – Old Irish alphabetic writing system

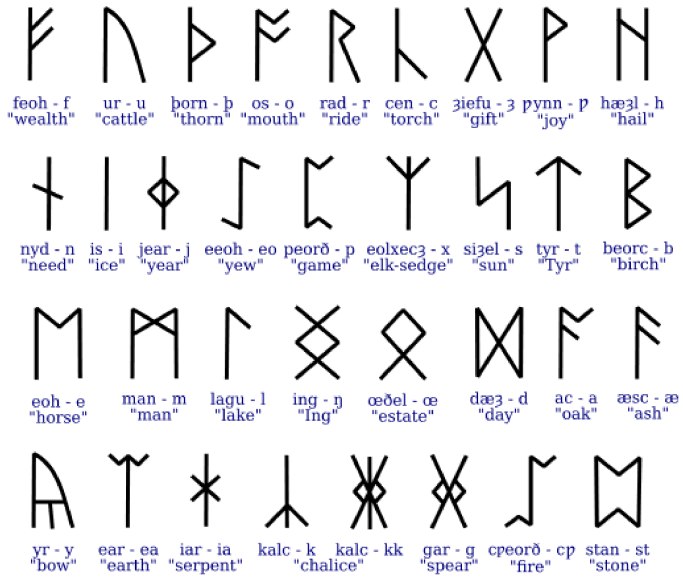


Retrieved from Sacred Chaotic Geometry on Tumblr

Around the 5th century, Anglo-Saxon settlers coming "from the area of north-west Germany and Denmark, and perhaps also the north-east of the Netherlands" (Hogg, 2002, p. 3) to Great Britain brought up their mixture of Germanic dialects. Two of the writing systems

prevailed, namely the runic alphabet (as known as *Anglo-Saxon Futhorc*, a variation of its sister-*genus* Scandinavian *Futhark*; see Figure 2), from the 5th to the 11th century, and later on the Latin alphabet with adaptations[2]. The Latin alphabet was set as the official writing system in English-speaking nations and it has continued on until present days. In addition, these Germanic dialects originated Ænglisc, or Old English, spoken not only in Great Britain but also in south Scotland.

Figure 2. Futhorc – Anglo-Saxon runic writing system

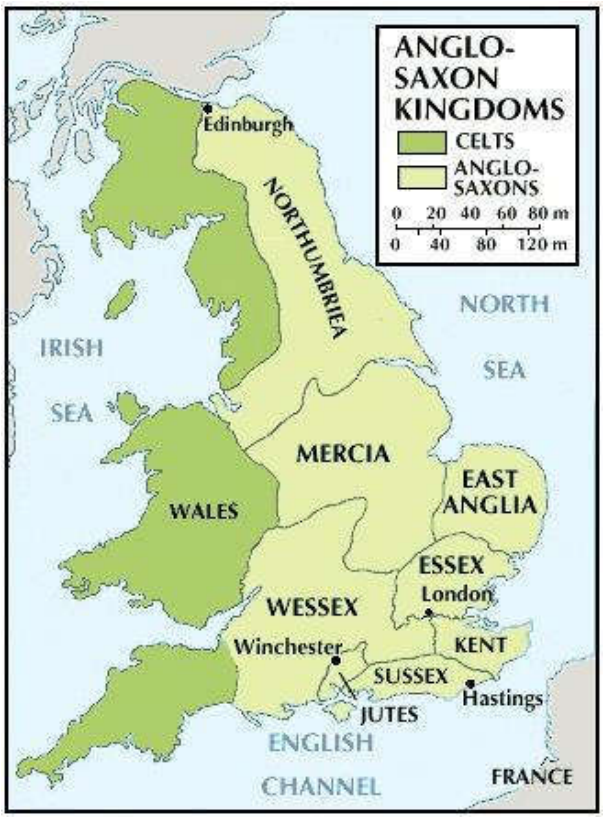


Retrieved from Wikipedia

At this point in history, England evolved geographically into seven kingdoms: Northumbria, Mercia, East Anglia, Essex, Kent, Sussex and Wessex – the latter the most powerful one, in the south, considering its military, economic, political and cultural influence over all the other kingdoms, as Figure 3 shows.

[2] Such adaptations are the letters þ, 'thorn', ð ('edh'), both nowadays standing for th; ȝ ('yogh'), nowadays represented by gh and the vowels æ, 'aesh' and œ, unnamed, with no specific graphic representation.

Figure 3. Map of England's seven kingdoms in the Middle Ages and Celtic territories



Retrieved from Silvia Lupi on Tes

Wessex had more resistance against the Viking invasions. Additionally, the first translations of the Bible from Latin to Old English took place there. This kingdom became a strong influential centre of the whole nation. So powerful that it called the attention of William I, the Duke of Normandy (in the north of France). This important character invaded Wessex in September of 1066 with his army of Normans. Then, he claimed the throne for himself one month later through the Battle of Hastings (note, in Figure 3, the geographical proximity between northern France and Hastings, in Sussex, where the battle took place), becoming known as William the Conqueror and William I of England. This historical event was then known as the Norman Conquest.

A new mark on the history of English arose, not only politically, economically and militarily, but also linguistically. Thanks to the Norman Conquest, English has suffered a dramatic change in its use and register, and therefore grammar. This mark led to the transitional period between Old to early Middle English, when a huge set of linguistic phenomena from Old and Norman French was introduced into English. For instance, the lexicon became full of Latin loanwords, which outnumbered Scandinavian loanwords. If one takes a look at Present-Day English and Present-Day French, or other daughter-languages of Latin, he or she will notice similarities. On the other hand, if one reads an Old English text such as *Beowulf* and is at least a bit familiar with German (*Hochdeutsch*), he or she would say, "This looks like German!" and would be barely able to read it properly without a translation or a gloss.

For Latin daughter-languages' speakers, like me, these claims seem reasonable. As a linguist, English fluent speaker and someone with an intermediate level of proficiency in German, I can notice that Old English indeed looks like German and other Germanic languages, like the Scandinavian ones, in the domain of morphology and lexicon. Naturally, I cannot say anything about similarities in pronunciation since it is impossible to retrieve data from spoken Old English to compare to its Germanic sister-languages. The same can be said about Middle English – naturally, no speech records are available, especially regarding spontaneous speech. Attempts in reconstructing the pronunciation of both past synchronies do exist, as Hogg (2002) does for Old English and Mossé (1968) for Middle English, for instance, as well as philologists when editing a manuscript. As in this research I am working with lyrics, these editors (cf. Wells, 1907; Hall, 1920; Brown, 1932 and 1939, and Duncan (1995), to cite just a few) make remarkable comments on the music behind these old manuscripts, citing other authors as well in their notes on each text. Despite all efforts, there is no way to retrieve real phonetic data from the dead speakers. Old texts is all we have.

Continuing the storytelling, Middle English was hugely influenced not only by French, but also Classical Latin, the brightest cultural language in Western Europe in the Middle Ages. Even after its death as a natural spoken language, its use as Classical Latin continued on especially in science, literature, law and religion. It is worthy to remember that at this point Christianity had reached Europe so strongly that it is almost impossible to set aside this fact when working with old manuscripts and their editions and interpretations. Being so, Classical Latin was the language of the Church in a period when most people could not read and write or understand the masses. As abovementioned, Wessex played an important role in this issue by giving place for the first translations of the Bible into Old English. In the next synchrony, Middle English, a huge amount of texts with words, expressions and even whole sentences and stanzas in Latin was not to be ignored. This kept growing as a tradition in Middle English. An interesting fact is that literate Englishmen were not the ones who could read and/or write in vernacular language, but in Latin, French *and* English. Clergymen writing religious lyrics used to write the same texts in three versions, one in Latin, a second in French and a third in English. It was even hard or impossible to know which was the translation of which; that is, *which one came first?* Perhaps not the one in English, for it had no social prestige, a framework enhanced after the Norman Conquest and the birth of Middle English with its French influence.

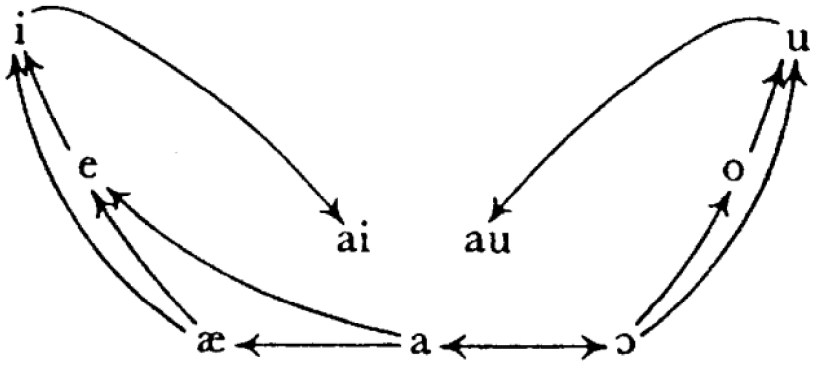
Speaking of the lack of social prestige of English, linguistic policy must be seriously regarded. *Why would a French-dominating English nation put efforts in promoting a non-prestigious language, even if it was their own?* – I ask. Considering the domain of morphology and spelling variation, Hogg assumes that "[…] such variation is not the result of error. […] As I said there, even a writer such as Aelfric, who took great care of the forms of his language, was not writing in a standard language. Such a type of language requires an educational and political infrastructure of a degree which, despite the undoubted

sophistication of the literate Anglo-Saxon community, was simply impossible. It is reasonable of talk about a focused language, that is to say, a range of variation of linguistic forms which a geographically defined literate community shared to consider degree, but without themselves imposing a well-defined set of spelling conventions, or by using some external source such as a national educational policy." (p. 22).

Prestige was conquered by English some centuries later. Until then, it kept changing, despite its lack of importance in Europe and "surprisingly" in England. If the 11th-12th centuries marked the transitional period between Old and early Middle English, the mid-15th century represented the transition between late Middle English to Early Modern English. The most significant fact about this latter period is the Great Vowel Shift, which is not the focus in this study, discovered by Otto Jespersen, who coined the term, in the beginning of the 20th century and explored by Giancarlo (2001).

Giancarlo (2001) brings on the revolution caused by the Great Vowel Shift from the late 14th to the 18th century. Middle English long vowels such as [i:] raise in place of articulation and turn into double vowels or diphthongs (p. 27), resulting in a transition from [i:] to [aɪ], for instance. An illustration of these changes is shown in Figure 4. This revolution was such in such a short timespan in the history of English that readers and learners of Middle English poetry find difficult to read aloud some words – while Middle English 'bite' was pronounced as 'beet' with a long [i:], the Early Modern English pronunciation is closer to Present-Day English 'bite', with the diphthong [aɪ]. Giancarlo's brief description does justice to Jespersen's (1909): "The great vowel shift consists in a general raising of all long vowels". A condition for the shift to take place was that the vowel must be not only long in duration, but also a stressed monophthong.

Figure 4. The Great Vowel Shift schema



Source: Giancarlo (2001:27)

The succeeding period of English is Modern English, here called Present-Day English.

Section 2 – Thesis objectives

Section 1 described briefly the history of English, essential to ground this study and update or inform the reader about the historical terminology concerning the history of the focused-upon language. Then, the next step is to state what the present study consists of, bringing on the main and specific objectives.

Firstly, this study aims to investigate the uses and meanings of the Middle English *haven*, focusing on its syntactic-semantic behaviour – the main objective. Then, it will lead us to discuss the verb valency of *haven*, "to have", as the one of the specific objectives, moving across linguistic changes (e.g. morphological changes and the process of grammaticalisation) in English over the covered centuries, namely the period between the 12th to the 15th centuries. The valency analysis will draw on Perini (2015). Finally, I will compare the studied centuries in a diachronic point of view and retrieve conclusions about changes and similarities.

Lastly, Section 3 will present the outline of this thesis.

Section 3 – Thesis outline: The structure of this study

So far, I have told a brief story of the English language, focusing on its main historical periods and transitions, showing important historical events such as the Norman Conquest and the Great Vowel Shift, in Section 1. Then, in Section 2, I have presented the general and main objectives of this study. Finally, yet importantly, the time to describe how this thesis will be organised has come.

This thesis will be divided in four Chapters, which, in turn, will have their own subdivisions. In Chapter 1, the theoretical framework delivered, showing two fields of Linguistics of our utmost interest here, namely Corpus Linguistics and Historical Linguistics. The former will draw on Biber (1993), Ghadessy, Henry and Roseberry (2001), Lawson (2001), Pustejovsky and Stubbs (2002), Gries (2009), McEnery and Hardie (2011) and Kübler and Zinsmeister (2015). The latter will rely on Bynon (1977/2012) and Shields (2012), moving forward to its interface with the former. Then, I will follow Kytö (2011) and Weisser (2015) with examples of historical corpora. Other interfaces will be shown as well (cf. Camargo, 2008, Rocha, 2014 and Cavalcante, 2015) with their respective corpora.

Chapter 2, accounting for the methodology, will have four subdivisions. The first one, Section 2.1, will describe the data and its collection process. As the data need to be balanced, this procedure will be handled in Section 2.2, relying on *R* software (R CORE, 2019). Then, as the data must be organised according to the variant spellings of *haven* (in the morphological domain) and classified according to its use (e.g. as a grammatical or lexical item), Section 2.3 will account for it. For this purpose, two spreadsheets, named S01 and S02 for these respective objectives, will be created and described. Spreadsheet S01 will be crucial

for a gloss-based analysis (following the Leipzig Glossing Rules, 2015) as well as for the spelling variation of the forms of *haven*. Spreadsheet S02 will prove essential to this Section for annotation will be done through the *UAM CorpusTool* software[3], version 3.3m, whose schema will be based on the use of each form of the focused-upon verb, namely grammatical or lexical. This software will be also useful to add the translations of the sentences containing any form of *haven* in the comments. These translations will be of my authorship except for the few available ones for some of the texts with adaptations when deemed necessary (cf. Delamar, 2003, Thomas, 2008 and Treharne et al., 2010).

Not only for balancing the corpus of study, but also for filtering and sampling the lexical occurrences of *haven*, which are of our utmost interest here, *R* software will be handy once again in Section 2.4. After the filtering process by means of a script to select only the ones marked as lexical, a sampling process will be conducted to sort 53 random occurrences. The analysis will be performed by investigating these sampled lexical occurrences.

After the corpus compilation and data processing, the results of the analysis on verb valency of *haven* taking place as a lexical verb will be presented in Chapter 3. Perini (2015) will ground the analysis with his schema of readings alongside the Middle English Dictionary (henceforth MED) developed by the University of Michigan (MI), which is corpus-based. Perini's schema for readings will be of great help in its adaptation considering the definition of the verb *haven* by MED and its senses and subsenses.

Chapter 3 will, in a first moment, deal with the results separately for each century covered (Section 3.1 for the 12th century, Section 3.2 for the 13th century, Section 3.3 for the 14th century and, lastly, Section 3.4 for the 15th century). Naturally, comments to ground brief comparisons will be made over these Sections, for, in a second moment, an in-depth comparison focusing on the readings to be conducted in the last Section (3.5). This, in turn,

[3] Available at: http://corpustool.com/

will follow a diachronic point of view by stating what has changed (and how) or prevailed during the Middle English period – in other words, from its first transition from Old to early Middle English until the second one, from late Middle English to Early Middle English. All procedures and results will be illustrated by means of Figures, Tables, Graphs and Lists.

In the end of this thesis, Chapter 4 will account for the Final Remarks, with two Sections. Section 4.1 will summarise the theoretical framework, methodology and results, whereas Section 4.2 will expose the difficulties and limitations of this study, mainly regarding the compilation of a Middle English corpus of lyrics and relevant works freely available to ground this research. In the end, ideas for further research will be provided.

CHAPTER 1 – THEORETICAL FRAMEWORK

This chapter aims to present, as its name suggests, the theoretical framework of this research, subdivided into two Sections, namely 1.1 and 1.2. Section 1.1 consists of the general literature on Corpus Linguistics, its potential interfaces with other fields of Linguistics, one of them Historical Linguistics, which concerns us here most. Considering the latter, I show and briefly describe examples of historical corpora for English and other languages (e.g. Icelandic and both Brazilian and European Portuguese), providing also their respective sources on the footnotes. As for Section 1.2, verb valency is discussed by calling the reader's attention to the origins of this terminology through a brief introduction of the notion of valency in Chemistry, then how this term came into Linguistics and what for. Since the specificity of this study draws on a syntactic-semantic analysis of the Middle English verb *haven*, "to have", which can assume divergent uses and meanings, for instance, behaving as a full-verb keeping its primary meaning of possession according to its etymology[4] and then expanding it throughout its history, or being employed as a grammatical verb in the past tense formation, verb valency must be discussed. More details on the uses and meanings will be provided in Chapters 2 and 3.

1.1. General literature: Corpus Linguistics in interface with Historical Linguistics

Due to Corpus Linguistics, it has become more feasible to analyse language or linguistic varieties for the most diverse purposes. McEnery and Hardie (2011) state that very well, pointing out that this field draws on a set of procedures to be followed by the linguist to achieve his or her goal on linguistic analysis. In some fields of linguistics, such as syntax, phonetics and historical linguistics, to cite just a few, Corpus Linguistics can be used but in

[4] The etymology of the focused-upon verb relies on the Online Etymology Dictionary, developed by Douglas Harper (2001-2019). Available at https://www.etymonline.com/search?q=have

different ways. Whereas in historical linguistics a corpus is the only source of data, in syntax and phonetics a corpus is *one* of the sources of data, being the intuitions of a native speaker preferable for most linguists, thus not exclusive. Back to historical linguistics, before the advent of corpus use new generations researchers relied on their intuition instead of practical methods to analyse one or more past synchronies of a language. Thus, the new empirical-based field opposes the long traditionally accepted approach based on a rationalistic basis, as we will see further in this very chapter. This seems to be a huge and impressive change indeed, since more and more data are becoming available as we speak, what results in a must to optimise the collection and processing of large bodies of data through technological advances. Kübler and Zinsmeister (2015) highlight the application of corpora for online consultation "for testing their linguistic hypothesis" (p. 3). Later I will present other uses for corpora, but for now I will attend to the definition of corpus, which is the core notion of Corpus Linguistics.

The definition for corpus has been exhaustively discussed among many researchers on the field of Corpus Linguistics (cf. Gries, 2009; McEnery and Hardie, 2011; Kübler and Zinsmeister, 2015; Weisser, 2015, among others). Based on these, although some divergences concerning the central definition of corpus may arise, a consensus is clearly observed in what concerns the need for a corpus to be a machine-readable collection of both written and/or spoken texts intended to be used in linguistic analysis. Such need contrasts with the assumption by Kübler & Zinsmeister (2015) that "[…] a long-standing tradition for working with corpora also in the pre-computer area, in particular historical linguistics, phonetics, and language acquisition" (p. 4). On the other hand, this need is essential in contemporaneity and has emerged due to technological advances. Naturally, given the increasing demand for large corpora in order to study rare linguistic phenomena, it seems more reasonable to compile linguistic data to be processed by computers, resulting in a faster and more assertive way to

make general assumptions about the focused-upon phenomenon. Many resources have become available to manipulate data according to the researcher's interest. Some of these resources are programming languages such as R and Python and softwares like *AntConc*[5] and *TextSTAT*[6] (useful, for example, for concordancing and the study of collocations), *SketchEngine*[7] (for instance, for text mining applications in translation studies), *UAM CorpusTool*[8] (useful for annotation schemata and additional comments), to cite just a few. This way, they can optimise corpus-based studies of language or linguistic varieties.

Given that a corpus must follow some criteria proposed by Corpus Linguistics, divergences may take place in this matter, and there is a very good reason for this fact. As Weisser (2009) assumes, "a corpus is [...] based on a specific set of design criteria influenced by its purpose and scope" (p. 13). For instance, if a linguist wants to investigate a very rare phenomenon in, say, spontaneous speech in Brazilian Portuguese, he or she can use a large corpus which represents this linguistic domain by having in mind that the larger the corpus, the higher the chances to find such phenomenon of his or her interest. For this particular purpose, size has a great importance, and it is naturally easier to access large bodies of data to search for this phenomenon. In contrast, if the linguist's purpose is to describe a not so rare or not yet well studied (or not yet studied at all) phenomenon in a given language or linguistic variety, a small corpus may indeed be useful and is by no means less important. For researches drawing on small corpora, see all sections in Ghadessy, Henry and Roseberry (2001), whose section 10, by Ann Lawson, proves worthy in an approach on parallel corpora. Irrespective of the type of approach, in order to study any linguistic phenomenon in the absence of an existing specific corpus for this purpose, in Lawson's own words, "[i]f no suitable corpus is available, the alternative is to make one" (p. 293).

---

[5] Available at: https://www.laurenceanthony.net/software/antconc/
[6] Available at: http://neon.niederlandistik.fu-berlin.de/en/textstat/
[7] Available at: https://www.sketchengine.eu/
[8] Available at: http://corpustool.com/

So far, I have briefly mentioned one of the central parameters of corpus design, namely size. As aforementioned, it can be strictly related to the researcher's purposes and needs. Additionally, representativeness and balance must be accurately introduced to this discussion for they are crucial notions in Corpus Linguistics. Representativeness and balance may be correlated as Biber (1993) suggests in his paper on corpus design. He calls the reader's attention to the decisions the researcher must make before designing a corpus, such as which population he or she intends to represent. In his own words, "[r]epresentativeness refers to the extent to which a sample includes a full range of variability in a population. In corpus design, variability can be considered from situational and from linguistic perspectives, and both of these are important in determining representativeness. Thus a corpus design can be evaluated for the extent to which it includes: (1) the range of text types in a language, and (2) the range of linguistic distributions in a language" (p. 243). Additionally, Biber's notion of population is related to a sampling frame. The data must represent the studied population, but it is unreasonable to retrieve and study all data from the population due to time and effort restrictions. Having this in mind, any corpus, irrespective of its size, must itself consist of a sample of the language or linguistic domain that is investigated. This sample must be balanced, especially in a corpus intending to represent different domains and genres of both written and spoken language. More details on the design of my corpus of study are addressed in Chapter 2.

Back to the purposes of the linguist, two questions arise: *Why design a corpus?* and *why use it?* As it was previously mentioned in this very chapter, the decisions on corpus design derive from the linguist's purpose. For instance, the present study aims to investigate the uses and meanings of the forms of the Middle English verb *haven* ("to have") in philological editions of lyrics dated from the 12th to the 15th centuries and, thus, a small corpus would be sufficient, because this linguistic domain (written production) and genre

(lyrics) are quite specific. Additionally, difficulties in having full access to an already-existing corpus of English lyrics dated from the 12th to the 15th centuries, also discussed in Chapter 2, had me make the decision to build my own corpus of study, which consists of a convenience sample. Given the research purposes, such corpus can be useful for the study of the occurrences of *haven* in this specific literary genre in the four mentioned centuries in medieval England.

As I previously mentioned, corpora can be applied to a large set of linguistic studies, enabling then interfaces between this field and others. The next paragraph will cover some examples of potential interfaces between Corpus Linguistics and other fields, followed then by the one which is of our utmost interest, namely Historical Linguistics, in a more detailed discussion.

One of these interfaces concerns Translation Studies, in which the linguist investigates language specificities in translation context, comprehending convergences and divergences between the source language and the target language (cf. Camargo 2008). Another interface is between Corpus Linguistics and Lexicography, whose aim is to study dictionarised terms in a language or linguistic variety, and its nature indicates metalanguage. For European Portuguese, I cite the CLP – *Corpus Lexicográfico do Português*[9], developed by the University of Aveiro, targeting the 16th to the 19th centuries. In addition, learner corpora can be the result of the interface between Corpus Linguistics and Pedagogy and Applied Linguistics, related to both spoken and written production by a second/foreign language learner (cf. Rocha, 2014). Corpus Linguistics can also be interfaced with Pragmatics. By means of the Language into Act Theory (Cresti, 2000), it is possible to study speech diaphasia, as does the C-ORAL Brasil Project[10] (Raso and Mello, 2009; Cavalcante,

[9] CLP – Corpus Lexicográfico do Português stands for "Lexicographical Corpus of Portuguese", my translation. Available at: http://clp.dlc.ua.pt/Inicio.aspx
[10] Available at: http://www.c-oral-brasil.org/

2015). Finally, an interface between Corpus Linguistics and Historical Linguistics is also possible. Since this study relies on a diachronic perspective, Historical Linguistics must also be taken into account.

Historical Linguistics "[…] seeks to investigate and describe the way in which languages change or maintain their structure during the course of time; its domain therefore is language in its diachronic aspect" (Bynon, 1977/2012, p. 1). Although this definition may be thought of as the fundamental one, this objective pointed out by Theodora Bynon may not be the main goal of Historical Linguistics, for, naturally, controversies may arise. In this approach, investigating how languages are formed, their origins and relatedness to each other (thus how they can be grouped together) is the main task faced by the historical linguist. In Shields' (2012) words, "[h]istorical linguistics constitute the sources of modern approaches to language study, especially through the efforts of the 19th-century Neogrammarians" (p. 1). He defends that this field has shaped the field nowadays called Linguistic Typology. Being so, this new field has originated from Historical Linguistics in his view, and can be used as a tool, proving its worthiness in how languages can be reconstructed and in the elaboration of principles concerning how languages evolve. According to Ramat (2012), linguistic typology means "the systematic cross-linguistic comparison that aims to discover the underlying universal properties of human language" (p. 1). This may be seen as a starting point for researching how languages change or maintain their structure, as Bynon stated.

In the pre-computer period, large bodies of data were starting to be compiled with data from many different languages. However, from the 17th to the 19th century, rationalistic views of language arise, setting aside the use of language itself, as well as the ones who use it. This opposes to the empirical basis of Corpus Linguistics. Only in the end on the 19th century the philosophical basis started to be dissociated from the linguistic basis, giving rise to the empirical method to analyse language.

An important consideration I must make is that reconstructing a language, in this case, a past synchrony of English, namely Middle English, is not necessary in this study since this linguistic period is already well documented, thanks to surviving written production, copied over and over, and their editions by philologists. Reconstructing a word's form and meaning, for instance, this research's focused-upon verb *haven*, is not among the scope of this study. This is so because this verb has entries in etymological dictionaries (e.g. *to have (v.)*, in the Online Etymology Dictionary (Harper, 2001-2019), presenting its origins from Old English *habban*[11], with its primary meaning of possession which then extended to others throughout the history of English). Additionally, the same applies for the dictionary adopted in this study, namely the Middle English Dictionary[12] developed by the University of Michigan (MI), whose entry is *haven*. Besides, reconstructing the potential phonetic/phonological representation of this verb is not my task either – but an etymologist's. Hence, by no means I intend to reconstruct historical speech, yet shed some light on how this verb changed and/or maintained its syntactic-semantic behaviour in the timespan from the 12th to the 15th century in written production, more specifically, lyrics.

Since this study uses Corpus Linguistics methods to compile linguistic data on philological editions of Middle English lyrics dated from the 12th to the 15th century for the purpose of investigating the uses and meanings of *haven*, it is possible to explore the interface between Corpus Linguistics and Historical Linguistics.

Kytö (2011) defends that "all historical linguistics is in a wide sense corpus-based", since the linguist must somehow collect linguistic data following certain criteria to study the history of a language or linguistic variety of interest. About using Corpus Linguistics' modern methods, she claims that it is unreasonable to nowadays, in the electronic era, to

[11] For more information on this verb's etymology, see https://www.etymonline.com/search?q=have
[12] Available at: https://quod.lib.umich.edu/m/middle-english-dictionary/dictionary

work with large bodies of data only manually. As previously stated in the discussion on Corpus Linguistics, the machine-readable nature of a corpus is a must in present times. When discussing the resources and methodologies for this type of study, Kytö goes back to the necessity of computerised data to analyse in a more optimised way data to substantiate an approach on linguistic change. As a sociolinguist, she calls attention to the urgency to process historical data automatically since this way "also facilitate[s] the statistical analysis of relationships between linguistic phenomena and linguistic or extralinguistic factors at work in language change". These extralinguistic factors would correspond to information about the writers, their profiles (gender/sex, age, level of literacy, social class, etc), the speech communities they were inserted in and so on, factors which may have indeed influenced the linguistic change. In addition, besides using corpora, there is no other option when studying a past synchrony of a language.

Some examples of historical corpora are, for English, the *Brooklyn-Geneva-Amsterdam-Helsinki Parsed Corpus of Old English*; the *Penn-Helsinki Parsed Corpus of Historical English*, composed by subcorpora of Middle English, Early Modern English and Modern British English, namely PPCMEP2 - *Penn-Helsinki Parsed Corpus of Middle English*, second edition, the PPCEME - *Penn-Helsinki Parsed Corpus of Early Modern English*, and the PPCMBE2 - *Penn Parsed Corpus of Modern British English*, second edition[13], respectively. These subcorpora of the Penn-Helsinki family served as a model to the PCMEP - *Parsed Corpus of Middle English Poetry*[14], still-developing by Zimmermann (2014-2019). About the Penn Treebank, which in the early 1990's targeted American English with a 4.5 million words, Marcus, Santorini and Marcinkiewicz (1993) reinforce the must for annotated large corpora to be consistently computerised, since this big corpora family aims

[13] Retrieved from https://www.ling.upenn.edu/hist-corpora/
[14] Available at https://pcmep.net/

PoS (Parts-of-Speech) tagging. The abovementioned parsed historical corpora follow this model and objectives for the sake of syntactic analysis.

Additionally, for Early Modern English, there is the CEEC - *Corpus of Early English Correspondence*[15] (Nevalainen et al., 1998-2019). This corpus family was started in 1993 targeting historical sociolinguistic data from letters covering approximately four centuries (ca. 1400-1800 AD) of the synchronic period of English known as Early Modern English, reaching 5.1 million words nowadays. Years later, especially after its completion in 1998, samples and extensions were taken from the original CEEC, giving birth to five daughter corpora[16]. In addition, several examples of historical corpora for English in both broad and specific domains (for instance, poetry and medicine) are provided and described by Martin Weisser in his website[17]. For other languages, I can cite the IcePaHC - *Icelandic Parsed Historical Corpus*[18], which, as its name suggests, is parsed and follows the model of the Penn historical corpora, targeting syntactic phenomena in Icelandic throughout its history from the 12th to the 21th century. For Portuguese, there is the *Tycho Brahe Parsed Corpus of Historical Portuguese*[19], also following the skeletal structure of the Penn family, covering written productions representative of the historical gap between the 14th and the 18th centuries. Furthermore, the Brazilian academic community has the *Corpus Compartilhado Diacrônico: cartas pessoais brasileiras*[20], which comes in hand to research the history of Brazilian Portuguese, whose consolidation process began in the 18th century, through written production, specifically personal letters, personal notes and plays from the 18th to the 20th century.

---

[15] Available at https://www.helsinki.fi/en/researchgroups/varieng/corpus-of-early-english-correspondence
[16] For more information on the daughter corpora of CEEC and the historical subperiods they cover, see https://www.helsinki.fi/en/researchgroups/varieng/corpus-of-early-english-correspondence
[17] Available at: http://martinweisser.org/corpora_site/historical_corpora.html
[18] For more information, see http://www.linguist.is/icelandic_treebank/Icelandic_Parsed_Historical_Corpus_(IcePaHC)
[19] Available at: http://www.tycho.iel.unicamp.br/~tycho/corpus/en/
[20] "Shared Diachronic Corpus: Brazilian personal letters", my translation. Available at: http://www.tycho.iel.unicamp.br/laborhistorico/corpora.html

In sum, the abovementioned historical corpora, among others for other languages and purposes (to study, say, syntax or semantics), proved themselves as useful tools for the study of past and present synchronies of the focused-upon languages or linguistic varieties. As noticeable, they are mostly parsed, relying on computational resources for their compilation process and parsing and, especially, on preliminary decisions on what population they are intended to represent and what they are used for. Then, I hope to have shed some light on the two important questions made above.

After the general literature on Corpus Linguistics was presented in the very beginning of this chapter, it was followed by potential interfaces between this and other fields. Moving forward, Section 1.2 will account for the specific literature to be used as the fundamental basis for the purpose of this research, in other words, concerning verb valency.

1.2. Specific literature: verb valency

If one claims to discuss what is called verb valency in Linguistics and for what this concept applies, its terminology must be introduced beforehand. The term "valency" is a jargon used in Chemistry, more specifically accounting for the Electronic Theory of Valency developed by Kossel and Lewis in the early 1900's[21]. According to this theory, which explains molecular formation through atomic combinations, valency is the amount of electrons an atom must "lose", "gain" or "share" in order to become stable, and this process requires completing the last layer of electrons in its electrosphere[22]. Then, each element has a number of slots in the last layer of electrons to be filled through a chemical bond with the

[21] In fact, Walther Kossel and Gilbert Lewis studied atomic structure separately in Germany and England, respectively. Both published their works in the same year, thus see Kossel (1916) and Lewis (1916).
[22] As explaining the Electronic Theory of Valency is under the domain of Chemistry and thus not under the scope of this research, for more information on the concept of valency in this field, see https://www.sciencehq.com/chemistry/electronic-theory-of-valency.html for a simpler explanation. For a deeper and Chemistry-based view of valency, cf. Partington (1921).

same or other elements. Depending on the atoms involved, different kinds of bonds may take place, with distinct elements. In other words, there are slots in the molecular structure that must be filled by forming chemical bonds with elements with certain chemical properties to reach the stability; depending on the valency of the elements involved in the bonds, the results may vary, with different combinations.

Back to Linguistics, In broad sense, this term regards the number of arguments that a predicate is able to control as crucial for the sentence to make sense. More than that, the arguments or complements are of less relevance given that verbs occur in many different contexts and this number varies. The syntactic-semantic functions of the complements are the key elements for defining a verb's valency.

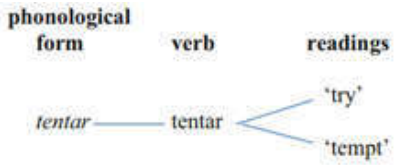
The French linguist Lucien Tesnière borrowed the term "valency" from the American polymath Charles Peirce, who significantly contributed to the field of Linguistics in what respects to pragmatism, semiotics and argument structure. For Peirce, verbs are compared to molecules since they can have elements attached to them forming different combinations for different purposes. This notion of valency from Chemistry into Linguistics was published in 1959 in a postumous work by Tesnière entitled *Éléments de Syntaxe Structurale* (Elements of Structural Syntax). His bright contribution is centered in the distinction between what the predicate requires as mandatory for the sentence structure, namely its arguments, and unrequired, optional structures, as known as adjuncts. This research will take this distinction into consideration, as I will show in Chapter 3. In a translation of Tesniére's work by Osborne and Kahane (2015), "[t]he semantics of a given verb tend to be predictive of the number of nouns with which it can combine" (p. xlvii). Tesnière does justice to Peirce in his rationale about comparing a verb to a molecule. In his own words, "[t]he verb may therefore be compared to a sort of atom, susceptible to attracting a greater or lesser number of actants, according to the number of bonds the verb has available to keep them as dependents. The

number of bonds a verb has constitutes what we call the verb's valency." (Book D, Chapter 97, §3, p. 239).

Based on the CMEL - *Corpus of Middle English Lyrics*, the syntactic-semantic approach for the description of the verb valency of *haven* will draw on Perini (2015). The author addresses the notion of construction, stating that a construction is sentence-sized for the purpose of analysing verb valency, although he recognises that other definitions (e.g. noun phrases) are also valid in the linguistic literature. Some verbs accept classifications, such constructions are called diatheses, and the "set of all diatheses of a verb" its "valency" (p. 5). Regarding each diathesis, Perini, in his study of Brazilian Portuguese, states that these diatheses include both a syntactic and a semantic component, namely: a) "an ordered sequence of form-class symbols […]", b) "one syntactic function […]", and c) "prepositions, individually identified […]", for the former, and "semantic roles of all NPs, AdjP, AdvPs, and prepositional phrases" for the latter (p. 6). In this forthcoming study, these steps will be followed in order to describe the verb valency of the Middle English verb *haven*.

As Perini (2015) assumes, valency dictionaries are made of a list of verbs in a language and their respective diatheses. Moreover, each entry in the dictionary must contain a phonological form, the verb itself, and then its different readings (or meanings). In order to classify verb readings, Perini (2015) suggests the schema presented in Figure 5 for the Brazilian Portuguese verb *tentar* ("to try/to tempt").

Figure 5. Perini's (2015) schema for the classification of the Portuguese verb *tentar* into its readings



Source: PERINI, 2015, p. 15.

As Figure 5 shows, the Portuguese verb *tentar* consists of the same phonetic and morphological unit. Then, two readings are provided, "to try" and "to tempt", representing each possible meaning of this verb. In other words, instead of two verbs with the same morpho-phonetic representation, there is only one with more than one meaning.

Van Gelderen (2011), from a cross-linguistic perspective of the valency changes in the history of English, claims that "[i]f languages differ in valency patterns, a historical linguist should be interested in which changes are possible" (p. 106). She also affirms, in contrast to authors like Nichols et al. (2004) and Comrie (2006), that there is no diachronic stability for English concerning verb valency, since many changes were observed for Middle English in comparison to other Germanic languages (p. 106). That being so, if changes in valency patterns can be observed cross-linguistically, it is also possible to investigate those changes throughout the history of a single language, as this study proposes for Middle English, namely the period from the 12th to 15th century.

Taking into consideration the fact that this study takes a diachronic perspective, CMEL's compilation process will be described in more details in Chapter 2. Then, after each step of the analysis, whose results will be discussed in Chapter 3, a comparison across the timespan from the 12th to 15th century will be conducted focusing on the syntactic-semantic

behaviour of *haven* taking into account its valency depending on which meaning it addresses as a full-verb.

CHAPTER 2 – METHODOLOGY

This chapter describes the corpus of study designed specifically for this research, namely the *Corpus of Middle English Lyrics* (CMEL), its compilation process as well as its metadata in Section 2.1. Having in mind the importance of balancing a corpus, Section 2.2 accounts for the application of this procedure, with the use of *R* software, version 3.5.3 (R CORE, 2019), to sample texts for each of the CMEL's subcorpora, namely 12th century, 13th century, 14th century and 15th century. The relevance of annotation, gloss and translation is discussed in Section 2.3. Section 2.4 discusses the process used to filter all occurrences of the forms of *haven* found in CMEL and sample the lexical ones.

2.1. CMEL description and data collection

The corpus of this study, named *Corpus of Middle English Lyrics* (CMEL), is a compilation of lyrics written in English dated from the 12th to 15th century. The choice of this timespan was due to the possibility to verify the uses and meanings of *haven* in lyrics from its intersection period between Old and Middle English (precisely the 12th century) to its end (mid-15th century, another transitional period between late Middle English to Early Modern English) and thus perform a comparison between them, focusing on what might have changed (see Chapter 3). Difficulties regarding the compilation and analysis of the CMEL are related to the lack of available material about philological editions of English historical literary works and scientific literature on the history of English due to copyrights and prices charged for access. Given the limited number of sources for the corpus compilation, a convenience sample was compiled following the criteria proposed by Corpus Linguistics.

CMEL compilation took into account the selection of lyrics establishing information about the century of each text as its primary criterion. Furthermore, considering that the data

of a corpus must be machine-readable (cf. Henry & Roseberry, 2002; McEnery & Hardie, 2011, and Kübler & Zinsmeister, 2015), after the texts selection, the following step was to process them in optical recognition and text edition softwares. Part of the available amount of texts was retrieved from printed books (cf. Wells, 1907; Hall, 1920; Brown, 1932, 1939; Mossé, 1968, and Duncan, 1995), whereas some were retrieved from the website of the *TEAMS* project of the University of Rochester[23] (NY) and the *Corpus of Middle English Prose and Verse* of the University of Michigan[24] (MI). Regarding the printed sources, the pages containing the texts selected were photographed and the image files were saved in .png format. Then, these files were read by *ABBYY FineReader14*[25] optical recognition software, version 14, which processed them and generated a .txt file as a result. These .txt files were manually revised in the text edition software *Notepad++* due to the occurrence of potential errors associated to the automatic nature recognition by the software. The texts retrieved online were copied and pasted in Notepad++, then saved as .txt files and cleaned with the use regular expressions[26], eliminating the numbers of the verses, spaces between the stanzas and between the lines. Afterwards, each .txt file was coded (see Table 1).

Table 1. Criteria to code the CMEL .txt files

Century	Century Code	Text Number	File Code
XII	A	001	A001
XIII	B	005	B005
XIV	C	037	C037
XV	D	001	D001

[23] Available at: https://d.lib.rochester.edu/teams/text-online

[24] Available at: https://quod.lib.umich.edu/c/cme/

[25] Available at: https://www.abbyy.com/

[26] A regular expression is "a formal method of specifying a text pattern" (Jargas, 2012, p. 19). The regular expressions used in this research, created by Rodrigo Araújo e Castro, were "^(0-9)+)?\t({7})?" (to remove numbers, tabs and spaces, in this order; " {4}", " {5}" (to remove extra four or five spaces in the beginning of the lines and in the middle of the verses); "^[0-9]+\r\n" (to remove numbers + empty lines), and "^\r\n" (to remove empty lines).

Table 1 shows the criteria to code the .txt files corresponding to their respective information about century (Letters A, B, C and D stand for the 12th, 13th, 14th and 15th centuries respectively) and text number instead of naming each file with the title attributed by their editors to each text. Most of them, lacking any proper title, were named by their editors with the first line, or even given new translated names, which were not changed. Therefore, one ought to consider that philological editions of the same manuscript may vary since the criteria adopted by their editors is most of the times not consensual. Moreover, other considerations account for the fact that there are more than one manuscript for the same text, for in the Middle Ages most of the written production was copied by many different hands over and over and passed through generations, so dialectal variation (which can consist of spelling variation and also scribes' errors to which corrections were proposed by the editors in order to fit the metrics better) may take place. In order to choose an edition as the most accurate, it had to be closer to the source manuscript, which must be the oldest one available, so that one can assume that the "same" text edited and said to represent a more recent century consists of an edition of a copy. Too modernised editions, targeting a lay audience instead of linguists were easier to find online, but were judged to not fully represent the written records of any of the centuries, or even not represent them at all, so they were avoided. The biggest problem to use such editions would be to rely on the loss of originality, what could indeed affect the spelling variation (e.g. the third person singular of *haven* is usually normalised as *hath*). Besides, the valency of *haven* in the original sentences would be affected as well.

Still considering the available editions of a given literary work, it is possible that one editor chooses to subdivide the work into parts, each with a name or simply "Part I", "Part II" and so on, whereas another decides to edit the text as a whole without subdivisions. This was the case for the texts representing the 12th century, naturally rarer and more fragmented; perhaps parts of the source text were lacking. The work called *The Proverbs of Alfred*, for

instance, edited by Hall (1920), the only source available, was subdivided into 23 parts named with Latin numbers (except for the first one, with the original Latin inscription "Incipiunt documenta Regis Aluredi", in PDE "There begin the documents of King Alfred", my translation), each one coded and saved separately since it was so in the used edition.

Finally, after the coding process, the .txt files were saved in folders, as Table 2 shows.

Table 2. Criteria to name the CMEL folders for each century

Century	Century Code	Folders	Content
XII	A	#1	Texts
		#2	Metadata
XIII	B	#1	Texts
		#2	Metadata
XIV	C	#1	Texts
		#2	Metadata
XV	D	#1	Texts
		#2	Metadata

The metadata for the CMEL texts were organised in .txt files, whose heading contained: a) text code (as detailed in Table 1); b) text name; c) status (*full* or *fragmented*); d) authorship; e) century/approximate date; f) subgenre/theme according to the editor's notes (e.g. *love lyrics*, a classification adopted by Duncan, 1995, or *Religious Lyrics: Dialogues between the Blessed Virgin and her Child*, as stated by Brown, 1939); g) number of lines (following the modern verse conventions adopted by the editors); h) number of types; i) number of tokens; j) type/token ratio; k) editor name; and l) references. In order to optimise the manipulation of the metadata this information was also organised in an electronic spreadsheet. As the texts were edited by philologists, the observations inserted after the heading in the .txt files of the metadata concerned exclusively of their authorship. Additional comments were included when necessary after the observations; for instance, the presence of verses written in other languages (as Latin) that were not removed since they were part of the

text in its entirety and crucial for its general meaning. As decided for the texts, the metadata were also saved in .txt format and coded as shown in Table 3.

Table 3. Criteria to code the CMEL metadata .txt files

Century	Century Code	Text Number	Metadata Code	File Code
XII	A	001	MD	A001MD
XIII	B	005	MD	B005MD
XIV	C	037	MD	C037MD
XV	D	001	MD	D001MD

As noticeable, the file codes were based on the same criteria decided for the .txt files corresponding to each text, but this time ending with "MD", which stands for "metadata".

Balancing a corpus is necessary in order to avoid bias (cf. McEnery & Hardie, 2011). For this purpose, a sampling process was required, which consisted of randomly selecting, through the *R* software, version 3.5.3, (R CORE, 2019), the amount of texts representing each century based on the century containing the lowest number of tokens (see Section 2.2). Afterwards, all the forms of *haven* in the sample underwent annotation through *UAM CorpusTool*, version 3.3m; in other words, this consisted of creating a new project and designing a schema to classify the uses of the forms of the focused-upon verb, namely into grammatical or lexical. Then, for each .txt file incorporated in the project, all the verb forms were marked and classified according to the proposed schema. For more details about this task, see Section 2.3.

2.2. CMEL size and balancing

When discussing small corpora, Lawson (2001) defends its dependency of a given context, in the case of this study, the specific literary genre lyrics written in Middle English and, centuries later, philologically edited by scholars. This research intends by no means to

represent either spoken or written Middle English in a general sense, or even describe its speech community, yet this specific genre was chosen to investigate the uses and meanings of the forms of the verb *haven* (and clearly CMEL could be further used to investigate other phenomena which can be found in this genre as well, as Section 4.2 in Chapter 4 will show). This specificity, alongside the reasons mentioned in Section 2.1 for a compilation of a convenience sample of philological editions of Middle English lyrics dated from the 12th to the 15th century, constitutes indeed a small corpus with specific purposes. If there is no such corpus available for use, in Lawson's words, "the alternative is to make one" (p. 293). Then, a balancing process is more than necessary. Besides, the decision to compile a small corpus relies not only in the scarcity of available material, but also in the ease to manipulate data context-based in a timely manner.

The necessity of a balanced corpus is discussed in McEnery & Hardie (2011). The authors raise a comparison between monitor corpora and balanced or sample corpora, the former consisting of corpora in which one can study language change over time and the latter trying to "represent a particular type of language over a specific span of time" (p. 8). Since this study aims to describe the uses and meanings of the verb *haven* in a specific past synchrony of English, Middle English (from the 12th to the 15th century) in a specific literary genre, namely lyrics, a sample corpus approach seems to be more adequate. Moreover, this view could enhance with Lawson's assumption, "The emphasis on size is of less relevance, however, for corpora for specific purposes, where quality is more important than quantity" (p. 293).

In raw numbers, CMEL reached 93,184 tokens, being 21,751 for the 12th century, 29,958 for the 13th century, 21,288 for the 14th century and 20,187 for the 15th century, so one can notice the unbalance between the number of tokens per century. A first attempt to count the types and tokens with *AntConc* software, version 3.5.7, was considered, but bugs were

observed, and this would undoubtedly affect the calculation of type/token ratio and the further sampling process. Then, the *R* software proved useful by means of scripts[27], one for each century, able to read the .txt files of all texts, count the number of tokens per text and its sum, the number of types per text and its sum and type/token ratio. Dataframes were run in order to turn possible to visualise the data. Finally, these were saved in spreadsheets in both .csv and .xmlx formats.

Being so, because the data representing the genre lyrics of the 15th century had the lowest number of tokens, 20,187, approximately 20,000, this subcorpus was used as the point of reference to balance the 12th, the 13th and the 14th centuries' subcorpora. Then, the *R* software (R CORE, 2019) proved useful once again in this balancing process by means of a script (see Appendices A1, A2 and A3) for sorting the texts based on a .csv spreadsheet containing the text number, text code and number of tokens per text. For the 12th century, in order to reach approximately 20,000 tokens, of the 32 texts, 4 were sorted to be set aside, in such a way that the sum of tokens of the 28 remaining texts reached 20,095. Likewise, for the 13th century, of the 71 texts, 46 were sampled to be disregarded, so the sum of tokens of the 25 remaining texts reached 20,148. Text B067, the largest one, with 10,934 tokens, could not be set aside since its removal would compromise the expected number of 20,000 tokens for the 13th century subcorpus. As for the 14th century, as the processing employed for the 12th and the 13th centuries, of the 54 texts, 3 were sorted to be removed, thus this sample consisted of 51 texts totalizing 20,313 tokens. Finally, CMEL's sample reached 80,743 tokens.

After balancing each of the subcorpora and thus CMEL, with each subcorpus containing approximately 20,000 tokens, the next step was to annotate all the forms of *haven*. This step will be more accurately described in Section 2.3.

[27] The scripts used here were adapted from Castro (2016).

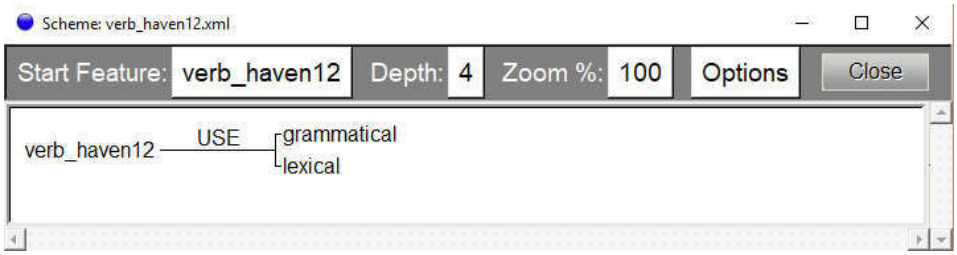
2.3. Annotation/gloss and translations (*UAM CorpusTool* and S01 and S02 spreadsheets)

According to Pustejovsky & Stubbs (2002), annotation comprises "[…] metadata that provides additional information about the text" (p. 1). The authors discuss teaching a computer how to process and provide the right data, but they are "[…] much less adept at understand language itself" (p. 2). As aforementioned, *UAM CorpusTool*, version 3.3m, has proved useful for marking all the forms of *haven* found in CMEL. The first step consisted of the creation of a new project for each century in .ct3 format *(verb_haven12*, *verb_haven13*, *verb_haven14*, and *verb_haven15*, respectively), set to manual annotation, whose schema accounted for the classification into lexical or grammatical use (see Figure 6). The second step was to add the folder corresponding to each century containing its respective sample (described in Section 2.2) and incorporate them to the project (see Figure 7). Finally, each text was manually annotated, in other words, all the forms of *haven* were marked in each text and classified based on the annotation schema (see Figure 8). As for the order of this task, it followed this sequence: the 12th century, the 13th century, then the 14th century and finally the 15th century. The third step consisted of the additional information about each occurrence, the translations of the sentences containing the forms of the verb in discussion. For some texts, translations proposed by authors (cf. Delamar, 2003 (for D031); Thomas, 2008 (for A027), and Treharne et al., 2010 (for A016)) were consulted, making changes on their solutions when judged necessary. However, most of the texts were translated freely by myself based on the editors' notes (if provided) and the Middle English Dictionary developed by the University of Michigan[28]. This dictionary was considered the most complete source available due to its information about the words' etymology, words' class, spelling variations and the enlisted uses and meanings, all evidenced by quotes retrieved from the texts of the *Corpus of*

[28] Available at: https://quod.lib.umich.edu/m/middle-english-dictionary/dictionary

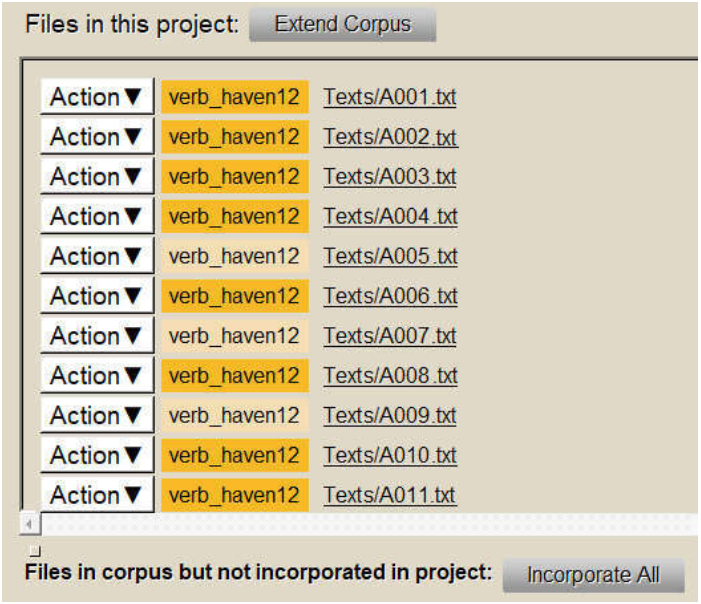
Middle English Prose and Verse. This reliability can be exemplified by the verb *haven* itself, whose entry presented thirteen uses and meanings also with collocates (see Appendix B). All this information proved crucial for the further analysis of its valency, as one will see in Chapter 3.

Figure 6. Annotation schema in UAM CorpusTool, version 3.3m



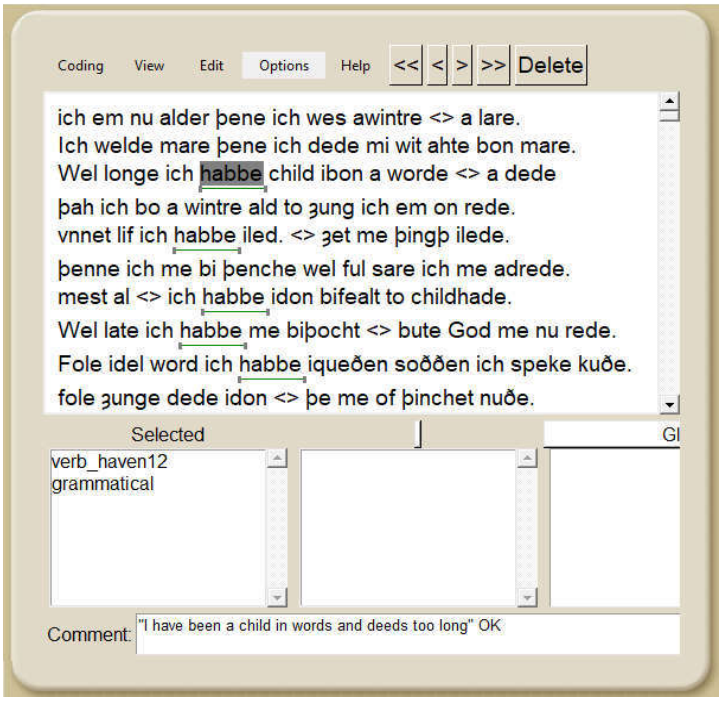
Retrieved from *UAM CorpusTool*, version 3.3m.

Figure 7. 12th century .txt files to be incorporated in UAM CorpusTool



Retrieved from *UAM CorpusTool*, version 3.3m, referring to all incorporated texts.

Figure 8. Marked forms of *haven* in Text A027



Retrieved from *UAM CorpusTool*, version 3.3m, referring to text A027.

Up to this point in the data treatment, the occurrences of the focused-upon verb into grammatical or lexical uses has not been discriminated. After translating all sentences where any form of this verb was attested, the spelling variation was taken into consideration and all forms were glossed in a spreadsheet type based on the Leipzig Glossing Rules (2015). This spreadsheet type, named "S01 - Gloss" (S stands for "spreadsheet"), consisted of information in columns about: a) tense, mode and negation, as known as "TMN", (H0 – infinitive; H0' - negative infinitive; H1 – present indicative; H1' - negative present indicative; H2 – past indicative; H2' - negative past indicative; H3 – present subjunctive; H3' negative present subjunctive; H4 – past subjunctive; H4' - negative past subjunctive; H5 – imperative; H5' - negative imperative; H6 – present participle, and H7 – past participle). Letter H stands for *haven*; b) gloss; c) forms of *haven*; d) total number of occurrences; e) texts (text code and how many occurrences per text); f) observations (for instance, in infinitive forms where the

conjugation took place in another verb (e.g. ‘mihte’, better translated to PDE as “could”) or when the subjunctive formation occurred with the employment of conjunctions such as variable forms of ‘if’ and ‘buten’ (“but”), both introducing a conditional clause, among other examples). In sum, S01 spreadsheets can be exemplified with Appendices C1, C2, C3 and C4, which report the existing forms corresponding to each gloss for each of the studies centuries. However, since this study focuses on this verb employed as a full-verb, a second spreadsheet was created, named S02 (see their representations in Appendices D1, D2, D3 and D4) concerning information about a) text code; b) forms of *haven* based on the order of the occurrences marked in each text in *UAM CorpusTool;* c) occurrence number (for instance, A007 - #1, #2); d) gloss based on S01; e) classification of its use into grammatical or lexical, and f) observations concerning the translations of the sentences where these forms took place as well as the ones regarding conjugation in other verbs syntactically related to any form of *haven*. Occurrences without clear classification were marked as “uncertain”, mostly due to uncertainty in the translations when the passages are fragmented, and thus disregarded. A low number of such occurrences took place, in such a way that their removal will not invalidate this study. Hence, Appendices D1, D2, D3 and D4 show the number of lexical and grammatical occurrences of *haven* per century.

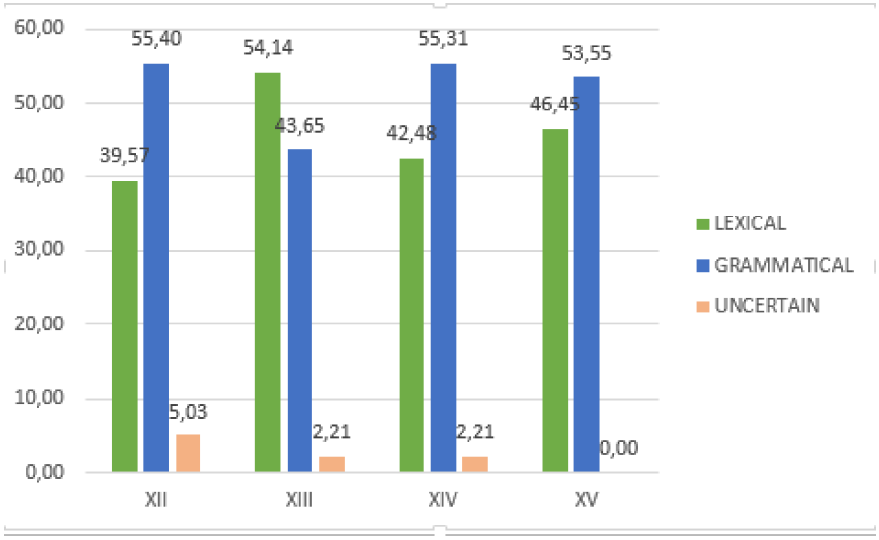
In sum, Table 4, followed by Graph 1, will represent the data shown in Table 7, Table 10, Table 13 and Table 16, in what respects to both the absolute and relative frequencies of the uses of *haven* per century.

Table 4. Absolute and relative frequencies of lexical, grammatical and uncertain occurrences of *haven* per century

Century	**Lexical**		**Grammatical**		**Uncertain**		**Total**	
	Abs.	**Rel.**	**Abs.**	**Rel.**	**Abs.**	**Rel.**	**Abs.**	**Rel.**
XII	55	39.57%	77	**55.40%**	7	5.03%	139	100%
XIII	98	**54.14%**	79	43.65%	4	2.21%	181	100%
XIV	96	42.48%	125	**55.31%**	5	2.21%	226	100%
XV	85	46.45%	98	**53.55%**	0	0%	183	100%

Graph 1 represents the relative frequencies of Table 4.

Graph 1. Frequency of items per century



Both Table 4 and Graph 1 intend to show a preview of the results. So far, they show that the highest relative frequencies of items in the 12th, 14th and 15th centuries (55.40%, 55.31% and 53.55%, respectively) represent the forms of *haven* employed as grammatical items. In contrast, the 13th century showed its highest relative frequency (54.14%) as corresponding to *haven* taking place as a full-verb. A hypothesis can rely on a skewing effect – the biggest text of this subcorpus is B067, with 10,934 tokens, which may have caused it. It will be further discussed in Chapter 3.

As aforementioned, there was not yet any discrimination between lexical and grammatical occurrences. However, based on Table 4, considering only the absolute frequencies of lexical occurrences of *haven* per century, which happens to be the focus of this research, the unbalance is evident. For this purpose, a sampling process of the referred occurrences is required, as explained in more details in Section 2.4.

2.4. Lexical occurrences sampling

This section presents the sampling process of the lexical occurrences found for each century in the sample retrieved from CMEL. This small corpus has already undertaken a balancing process for the number of tokens for each subcorpus (see Section 2.2), and now it is time to perform another sampling. In order to fulfill this task, the number of lexical occurrences of *haven* for the 12th century was taken as a reference (55) since it is the lowest, so that this process was employed precisely for the other three centuries. Then, for each the 13th, 14th and 15th centuries a .csv spreadsheet containing a) text code; b) occurrence number (per text); c) occurrence ID (general), and d) classification (into lexical or grammatical or uncertain) was created. It was then imported to *R* software and a script was run for the purpose of filtering only the lexical occurrences and then sorting 55 of them to be analysed (see Appendices E1, E2, and E3), generating another .csv file as a result.

A third spreadsheet, named S03 was created for each century precisely for the morphosyntactic gloss-based (following Leipzig Glossing Rules, 2015) analysis of the sentences where the sampled lexical forms of *haven* took place, as Table 5 shows.

Table 5. Example of a glossed sentence from the 12th century's sample

A014	**Occ01**	Tier 1	ME text	***if***	***þu***	***hauest***	***seorewe***
		Tier 2	Gloss	if-CONJ	you[2SG.SBJ]	have-PRS.2SG.SBJV	sadness[SG.DIROBJ]
		Tier 3	Reading			4b(a) 'to have (sadness)'	
		Tier 4	Arguments		Arg1		Arg2
		Tier 5	Translation	*"If you have sadness, […]"*			

For each occurrence, a schema was organised in Table 5 as follows: a) text code (e.g. A012); b) number of occurrence (in the order of the text, e.g. Occ01); c) tiers – Tier 1: ME text (each word in a cell); Tier 2 – Gloss (in each word's corresponding cells); Tier 3 –

Reading (based on the Middle English Dictionary developed by the University of Michigan); Tier 4 – Arguments (of the respective form of *haven*), and Tier 5 – Translation (of the whole sentence in order to provide the context of use). Tier 3 was given great attention since the readings were codified the same way their senses and subsenses were found in the MED (e.g. 'to have (sadness)', code *4b(a)*). Likewise, Tier 4 was important for the further analysis of the verb valency of the focused-upon verb (see Chapter 3). This procedure was carried out first for the 12th century and then to the others.

When designing S03 for the 13th century, two classification errors were found, so that the reference number of tokens had to be changed to 53. The S03 for the 12th century was already complete and a new sampling process would delay the analysis since it could result in a different sort and then should be remade. Being so, in order to spare time and effort I used a command line[29] on *R* to sort one occurrence to be disregarded from the 12th century's S03, since this very occurrence took place in one of the texts previously sorted to be disregarded, whereas the other texts sorted for this purpose did not contain any form of *haven* so they did not appear in the S03 for the 12th century. As a result, for this century, the 34th occurrence in order of apparition was marked to be set aside. The next step was to adapt the *R* scripts for both the 14th and 15th centuries to sort 53 lexical occurrences instead of 55 (this is why Appendices E2 and E3 differ from Appendix E1). Finally, their respective S03s were designed following the one proposed for the 12th and 13th centuries.

Even though all spreadsheets designed for this research will prove themselves useful in Chapter 3, the S03s for all centuries will play a very important role in the results, perhaps being of greatest importance. This is so because the S03s for each century consist of the sampled lexical occurrences (which as previously mentioned is of our utmost interest here)

[29] The command line, run on R, was: sample(1:55,1). That is, sampling 1 number from the set containing the numbers 1 to 55.

where all the information that could lead to the investigation of the verb valency for each century is presented.

Before introducing the results through Chapter 3, I will explain how the analysis of the verb valency of the sampled lexical occurrences of *haven* for each of the centuries will take place.

Firstly, let us consider S03. All the readings found will be shown through a first set of figures following Perini's (2015) schema (see Figure 5 on Chapter 1), namely Figures 8, 9, 10 and 11, respectively for the 12th, 13th, 14th and 15th centuries. These readings will be coded according to MED and numerically ordered, and then they will be discussed in each Section accounting for each century. Afterwards, each of the readings will be followed by their respective valency and diatheses, which will be exposed through a second set of figures, namely Appendices F1-A and F1-B (for the 12th century), I1-A and I1-B (for the 13th century), L1-A and L1-B (for the 14th century) and O1-A and O1-B (for the 15th century). This is essential to show, for each century, the readings and valency found in the samples. An important consideration about this matter lies on word order flexibility that still took place in Middle English, resulting in the allowance of all the six possible orders of Subject, Verb and Object(s). Mossé (1968) states that very well; in his own words, "[d]espite the progressive impoverishment of flexion the word order in ME in which the principal elements of the sentence, subject (S), verb (V), object (O) and adjuncts (A) were placed was still very flexible […]" (p. 122)[30]. Also, he provides examples of all six syntactic orders (cf. Mossé, 1968, p. 122). In addition, one must have in mind that in the chosen genre for this research, lyrics, variation is frequent, especially in syntactic order due to metric reasons – as those texts were

[30] Although Mossé (1968) addresses the adjuncts in his discussion about syntactic flexibility in ME, this category will not be considered in this research for they are not mandatory in the sentence. Mandatory elements are, here, *haven*'s arguments.

meant to be sung or read aloud, such adaptations would be very well appreciated. Thus, one might expect a larger set of argument structures in CMEL's data.

Secondly, once the readings and valency for each century are presented and discussed in their respective section, the data found in the S01s will be further discussed through tables showing the most frequent forms for each gloss. Although in some studies researchers look for rare occurrences into hapax legomena, in this research they will not be regarded due to time restrictions, except when it is the only occurrence for a given gloss or when two forms occurring only once overlap for the same gloss. In these cases, notations in the tables will be added. Apart from that, the most frequent forms in the tables will be highlighted and a brief clarification will be conducted.

Thirdly, the S02s will present both the absolute and relative frequencies for each classification (grammatical, lexical or uncertain) in each century. Considering all spreadsheets mentioned, a list containing the most frequent readings codified according to MED's senses and subsenses will be provided alongside their number of occurrences per century. In order to decide a cutpoint, the median was then calculated based on this list. For the 12th and 15th centuries, the numbers were both 5.5 (rounded to 5); for the 13th century, it reached 5; for the 14th, it surprisingly resulted in 4. Since the 15th century contained a more unbalanced distribution of occurrences and a greater number of hapax legomena, the cutpoint for all centuries was downsized to 4 so that more readings could be discussed for the 15th century as for the others.

Then, the results for each century will be discussed in Chapter 3 followed by a cross-century comparison.

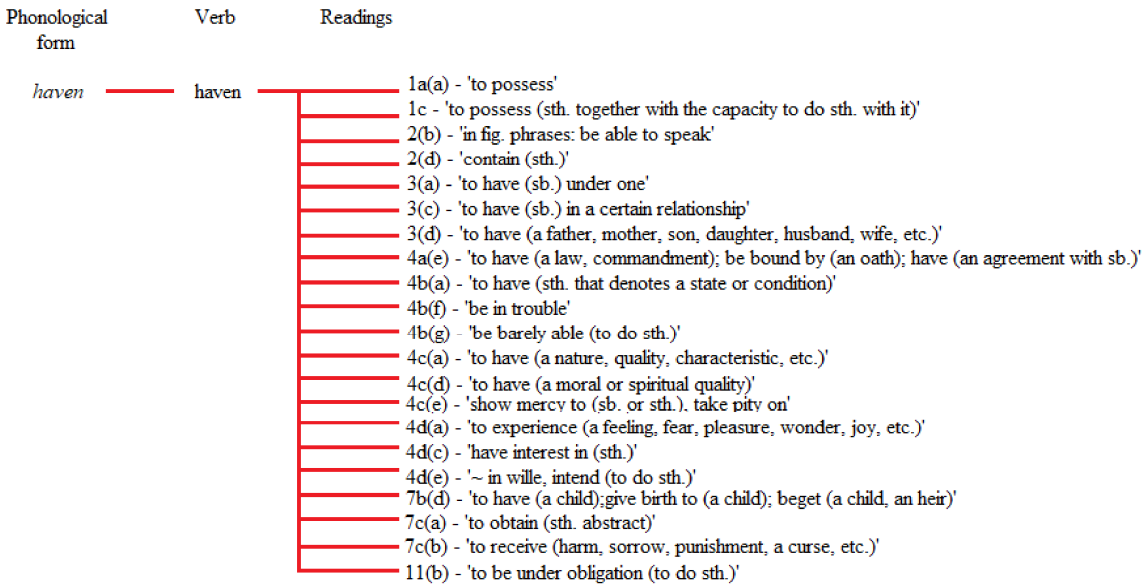
CHAPTER 3: RESULTS

This Chapter will bring on the results for each century through Sections 3.1, 3.2, 3.3 and 3.4 respectively. In each of them, I present the readings found in the sample, then explain the valency for each and discuss their diatheses. Then, the data extracted from the spreadsheets are shown through tables, figures, lists and graphs and discussed. In the end, in Section 3.5, a comparison between the data for the four centuries of interest will be handled.

3.1. 12th Century

I begin by introducing the readings retrieved from the Middle English Dictionary (MED; see Appendix B) developed by the University of Michigan found in the sample of the lexical occurrences of *haven*. These readings are coded, numerically ordered, and followed by their respective senses and subsenses, which will be called readings. The schema used by Perini (2015) in Figure 5 is adapted to show these readings as Figure 9 illustrates.

Figure 9. Readings based in MED found in the sampled lexical occurrences for the 12th century



As noticeable in Figure 9, 21 different readings were attested. Then, based on that, the valency for each reading will be presented (see Appendices F1-A and F1-B) by means of a second set of figures as an adaptation of the schema in Herbst et al. (2004) in their valency dictionary for Present-Day English (PDE)[31]. Not only the valency will be presented, but also examples extracted from this subcorpus in List 1 (See Appendix G).

Appendices F1-A and F1-B alongside Appendix G show each reading found in the 12th century's sample of lexical occurrences. Herbst et al.'s schema for a PDE valency dictionary, which is also corpus-based, at least for the verb *have* seems to be more concise, whereas the one I am hereby presenting is more detailed according to sampled data from a small corpus, whose readings classification system comes from MED. Even though I make use of a PDE valency dictionary to serve as a model, by no means this study has amongst one of its purposes any comparison between ME and PDE. The choice for a PDE valency dictionary relied on the lack of any valency dictionary available for ME.

After this note, let us comment the valency attested for the 12th century. The first observation is that no form of *haven* from the sample of this subcorpus took place in the passive voice. A second one is the occurrence of relative clauses, in which the relative particle is duplicated (e.g. Arg1 followed by Arg1').

List 1 (Appendix G) does not provide examples for all the argument structures shown in Appendices F1-A and F1-B, which would thus consist not of examples but the whole set of sentences analysed in the sample. Its purpose is simply to exemplify the occurrences. Even so, List 1 shows that: 1) *haven*'s diathesis "[N]" requires a Subject (Arg1) and a Direct Object (Arg2); 2) "[N PP]" requires a Subject (Arg1), a Direct Object (Arg2) and an Indirect or an Oblique Object (Arg3), and 3) infinitive clauses can consist of one of *haven*'s arguments, taking place alongside verbal particles such as 'to', 'for' and the combination 'for

[31] In the absence of a valency dictionary for Middle English, Herbst et al. (2004) was chosen as a model.

to' or simply by means of the morphological ending *–en* (and its spelling variants), as well as, in some cases, be inferred by the context. When the infinitive clause has *haven* as its main verb, it is highlighted in the diathesis as "V(haven)" or "V(to-INF(haven))", for instance. The purpose was to differentiate between infinitive clauses with other verbs and the ones with *haven* itself.

Now that the valency of *haven* in the 12th century's sample were presented and discussed, S01 and S02 deserve some attention. Table 6 represents S01 and contains only the most frequent forms of the focused-upon verb by gloss.

Table 6. S01 for the 12th century: Most frequent forms of *haven*

TMN	Gloss	Forms of *haven*	#Occurrences	Text distribution
H0	haven-INF	habbenn	8	A030
		habben	5	A027
H0'	NEG-haven-INF	naffde	1 (only)	A030
H1	haven-PRS.1SG.IND	habbe	18	A027 [7]; A030 [10]; A032 [1]
	haven-PRS.2SG.IND	hafesst	1 (only)	A030
	haven-PRS.3SG.IND	haueþ	8	A009 [1]; A012 [1]; A013 [2]; A016 [1]; A018 [1]; A019 [1]; A027 [1]
		haueð	6	A027
	haven-PRS.1PL.IND	hafenn	2	A032
	haven-PRS.3PL.IND	habbeð	6	A027
H1'	NEG-haven-PRS.3SG.IND	nafð	1 (only)	A027
	NEG-haven-PRS.3PL.IND	nabbeð	2	A027
H2	haven-PST.2SG.IND	heuedest	1 (only)	A016
	haven-PST.3SG.IND	haffde	24	A029 [5]; A030 [19]
	haven-PST.3PL.IND	haffdenn	13	A030 [12]; A032 [1]
H2'	NEG-haven-PST.3SG.IND	naffde	2	A030
	NEG-haven-PST.3PL.IND	naffdenn	2	A030
H3	haven-PRS.2SG.SBJV	hauest	2	A014
	haven-PRS.3SG.SBJV	*haueþ*	*1*	*A017*
		habben	*1*	*A027*
	haven-PRS.1PL.SBJV	habbe	1 (only)	A027
	haven-PRS.3PL.SBJV	habben	1 (only)	A025
H3'	NEG-haven-PRS.3SG.SBJV	nabbe	1 (only)	A022
H4	haven-PST.1SG.SBJV	hefde	1	A027
	haven-PST.3SG.SBJV	*hefð*	*1*	*A027*
		hefde	*1*	*A027*
	haven-PST.1PL.SBJV	hefden	1 (only)	A027

In order to research the forms of *haven* in CMEL, all the forms must be beforehand regarded in its orthographic character. As spelling variation tended to occur in written production in the Antiquity and in the Middle Ages (the latter is of utmost interest here), it is expected that the verb *haven* would present it too. However, it is also expected that some spellings have preference over others. To exemplify, Table 6 shows the gloss coded as "H0", namely "haven-INF", for which two forms are preferred – *habbenn* and *habben*, respectively with 8 and 5 occurrences. If one pays attention to the last column, he or she might conclude that each of these two forms were preferred in a single text each. This leads to the the question *why does, say, "habbenn" occur only in text A030?* and the most reasonable answer might be "because the author of this text was rather familiar with this spelling than with its alternative ones".

On the other hand, for the gloss coded as “H1”, as known as “haven-PRS.1SG.IND”, the preferred spelling is *habbe*. No question needs to be asked in a first moment about the 18 times this particular form takes place in the 12th century subcorpus; the explanation relies on the text distribution – these 18 occurrences of *habbe* took place in more than one text. However, the question above may rise once more, this time asking why, amongst these 18 occurrences, 10 belong to text A030, 7 to text A027 and only one to text A032. To deepen the answer, one might say, after checking the data, that this is due to the fact that A030 is a very long text (13,507 tokens) in comparison to A032 (1,322 tokens). The answer makes sense, since the chance for a given spelling to be preferred over others might be higher in a text with more tokens, adding the author’s preference and knowledge of or familiarity with the writing system of the language.

Considering the glosses, two contrasts deserve attention. The first one is related to negation – negative forms are rarely used, whereas non-negative ones take place quite often. The second one refers to the distinction between realis and irrealis – indicative forms are more common than subjunctive ones. Both contrasts lead us to a morphology-based discussion. In a brief comparison with Old English, Mossé (1968) states that “[w]hile Old English is still a richly inflected speech, Middle English, by way of simplification, is a speech poor in inflection. The tendency was hastened by the social upheaval and the mixture of tongues that the Norman Conquest brought about, but it was already noticeable in Old English from the end of the 10th century” (p. 44). The author shows morphological changes suffered by Middle English, focusing on the declension system, which happens to be very strong in Old English and is impoverished in Middle English. As a matter of fact, the mixture of tongues he refers to is highly influenced by Old French since this language brought by William the Conqueror and the Normans under his command became privileged, what caused English to be of no privilege in many social environments. Even before the Norman

Conquest, also Latin was a privileged language of culture in Western Europe, what explains words, expressions and even whole sentences in Classical Latin in the lyrics present in CMEL, especially in religious written production. Scandinavian loanwords were also common, but still Old French (Parisian dialect) has had the highest prestige in the conquered England for the centuries to come.

Back to the less usage of negative and irrealis (subjunctive) forms, it might have to do with the French influence. Not only most declension is lost, but also the verbal inflection system is impoverished. Separate words with their own meaning and/or grammatical purposes come in hand to express negation, such as the French double negation system. In addition, words such as "if" and "when" to express irrealis (subjunctive) also arise.

Still regarding the unsampled forms of *haven* and keeping the flow of the discussion in the broad sense of linguistic change, grammaticalisation must be regarded. This research has no aim to exhaustively explain what this process consists of (cf. Kurylowicz, 1965/1975; Bybee et al., 1994; Pagliuca, 1994; van Gelderen (2005) and Eckardt (2006), among others). Being so, what can be very briefly and superficiously said about grammaticalisation is that it consists of the gradual loss of lexical context of a given word when/and it begins to be employed as a grammatical item (for OE *habban*, to express past tense, which continued on until PDE). In the data retrieved from CMEL, all forms of *haven* had to undergo a classification process into grammatical, lexical and uncertain occurrences, as explained in more detail in Chapter 2. Table 7 shows this classification below.

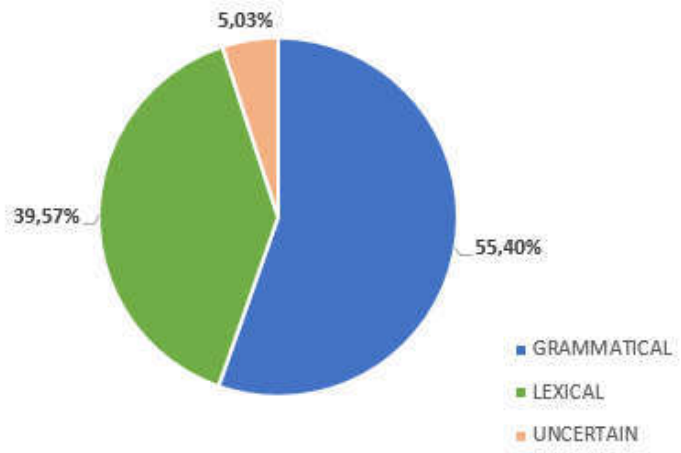
Table 7. S02 for the 12th century: Classification of the forms of *haven* (unsampled) and their respective absolute and relative frequencies

Grammatical		**Lexical**		**Uncertain**		**Total**	
Abs.	**Rel.**	**Abs.**	**Rel.**	**Abs.**	**Rel.**	**Abs.**	**Rel.**
77	55.40%	55	39.57%	7	5.03%	139	100%

The information shown in Table 7 account for the S02 for the 12th century. As mentioned in Chapter 2, this spreadsheet refers to the abovementioned classification into grammatical, lexical and uncertain, each containing their respective absolute and relative frequencies. The methodology adopted in this research has clarified the filtering and sampling processing of these occurrences by means of two scripts on *R*. Special attention is given to the predominance of grammatical occurrences (55.40%) in this subcorpus. This fact points to the use of the focused-upon verb as a grammatical item in this transitional period between Old and Middle English. Grammaticalisation was already taking place in Old English, and so the process of a word gradually losing its primary meaning (in the case of *haven*, which derives from Old English *habban* meaning "to own, possess") continued on. Table 7 makes sure this tendency.

Below, Graph 2 illustrates what is presented in Table 7.

Graph 2. S02 for the 12th century: Relative frequencies according to their classification



In respect to Graph 2, different colours show the three abovementioned categories. Blue stands for "grammatical", green for "lexical" and pink for "uncertain". Having this in

mind, blue prevails (55.40%), followed by green (39.57%). The uncertain occurrences, in pink, as previously stated, are set aside.

Now that the valency found in the 12th century's subcorpus were shown and described in two sets of figures (see Appendices F1-A and F1-B), as well as the most frequent forms of *haven* according to their glosses (Table 6), followed by the classification of all occurrences of *haven* in the focused-upon century into categories of use, it is the time to go further. Back to the readings of *haven* according to Perini's (2015) schema following the senses and subsenses from the MED, here called "readings" to follow Perini (2015), a list will be provided with a different type of organisation (see Appendix H for List 2).

List 2's (Appendix H) ordering relied on the frequency of use of each reading found in the 12th century's sampled lexical occurrences. The most frequent are *7c(a) - 'to obtain (sth. abstract)'* (10 occurrences), followed by *1a(a) - 'to possess'* (8 occurrences); *4d(a) - 'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'* (5 occurrences); *3(c) - 'to have (sb.) in a certain relationship'* (4 occurrences), and *4b(a) - 'to have (sth. that denotes a state or condition)'* (4 occurrences), as exemplified below. Note that the primary meaning of possession is surprisingly the second most frequent, not the first, as expected, especially considering that the 12th century is the transitional period between Old and Middle English.

1a(a) – 'to possess'
[N] – *ech mon mid he* ***hauet*** *mei buggen houene riche*
each man with he has may buy heaven kingdom
"each man is able to buy the kingdom of heaven with all he has" - Arg1 V Arg2

4d(a) – 'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'
[N] – *ach þo þre habbeþ scome and grome and oft fele sorȝe*
but the others have shame and wrath and other great sorrows
"But others have shame and wrath and other great sorrows" - Arg1 V Arg2

[N PP] – *whatt menn mihhtenn habenn niþ ne wraþþe ȝæn heore owwþerr*
what men could have envy neither wrath against each other
"what men might have neither envy nor wrath against each other" - Arg1 V Arg2 Arg3

3(c) – ‘to have (sb.) in a certain relationship’
[N] – *freond þat he habbe*
friend that he has
"a friend that he has" - Arg2 Arg2’ Arg1 V

4b(a) – ‘to have (sth. that denotes a state or condition)’
[N] – *nabbeð hi nane blisse*
not_have they no peace
"[they both (cold and heat) do them enough suffering;] they have no peace" - V Arg1 Arg2

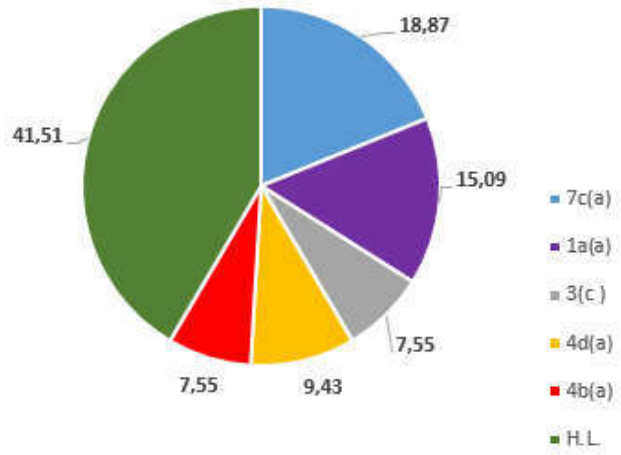
[N PP] – *of hete hi habbeð misse*
of heat they have need
"[when they come again to cold,] they have need of heat" - Arg3 Arg1 V Arg2

In this sample, 11 readings were attested as hapax legomena. Since the number of occurrences greater or equal to 4 was considered the cutpoint to set the boundaries between the most relevant occurrences and the other, all the other occurrences below this number were grouped with the hapax legomena (H.L.), totalising 22 occurrences. As a result, Table 8 and Graph 3 show the absolute and relative frequencies of each most frequent reading.

Table 8. S03 for the 12th century: Most frequent readings according to MED (sampled lexical occurrences)

Reading code	Reading meaning	Abs. Freq.	Rel. Freq.
7c(a)	‘to obtain (sth. abstract)’	10	18.87
1a(a)	‘to possess’	8	15.09
4d(a)	‘to experience (a feeling, fear, pleasure, wonder, joy etc)’	5	9.43
3(c)	'to have (sb.) in a certain relationship'	4	7.55
4b(a)	‘to have (sth. that denotes a state or condition)’	4	7.55
H.L.	Other occurrences	22	41.51
TOTAL		**53**	**100%**

Graph 3. S03 for the 12th century: Most frequent readings according to MED (sampled lexical occurrences)



In Graph 3, dark green stands for the hapax legomena group (41.51%); light blue represents *7c(a) - 'to obtain (sth. abstract)'* (18.87%); purple points to *1a(a) - 'to possess'*; orange depicts *4d(a) - 'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'* (9.43%); grey denotes *3(c) - 'to have (sb.) in a certain relationship'* (7.55%) – same relative frequency as *4b(a) - 'to have (sth. that denotes a state or condition)'*, expressed by red.

The following Section will account for the results referring to the 13th century's subcorpus.

3.2. 13th Century

For Sections 3.2. 3.3 and 3.4 (accounting for the 13th, 14th and 15th centuries respectively) the main structure adopted for Section 3.1 (in respect to the12th century) will be followed. This Section will present and explain the results found for the 13th century's subcorpus. I bring forward that there is a hypothesis to be introduced and discussed. First, I will present in Figure 10 the readings based on the senses and subsenses of *haven* according to the Middle English Dictionary (MED; see Appendix B).

Figure 10. Readings based in MED found in the sampled lexical occurrences for the 13th century

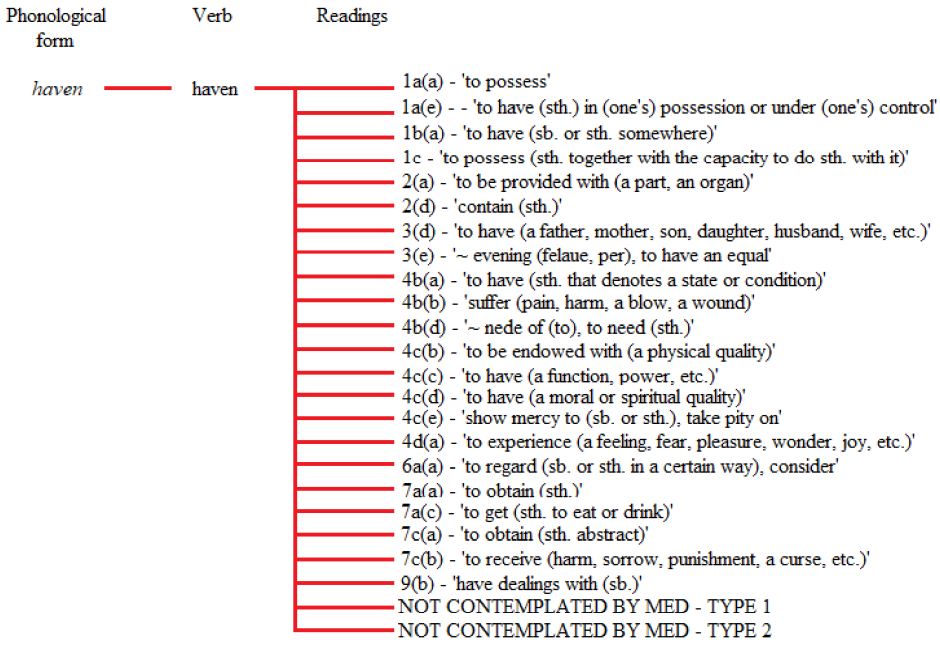


Figure 10 shows all the readings found in the 13th century's sampled lexical occurrences. As previously explained, they are organised in numerical order considering the codes from MED, followed by their senses and subsenses as known as readings in this study. However, two readings are not contemplated by MED, here called "Type 1" and "Type 2". Type 1 stands for "to be quiet" and Type 2 stands for "to have an answer". Even following the dictionary of interest, awareness of the fact that some occurrences might not be classified according to it is crucial, and for this reason, I was inclined to name these readings as "NOT CONTEMPLATED BY MED – TYPE 1" and "NOT CONTEMPLATED BY MED – TYPE 2". 24 readings, including the latter ones, were attested in total.

As for both Type 1 and Type 2 readings, examples follow.

NOT CONTEMPLATED BY MED – TYPE 1 – 'to be quiet'
[N PP] – *habbe he is tunge under gore*
have he his tongue under clothing
"[For when he has done his deed] he shall have his tongue under clothing"

NOT CONTEMPLATED BY MED – TYPE 2 – 'to have an answer'
[N] – *an ȝet ich habbe an oþer andsware*
and yet I have another answer
"and yet I have another answer"

Based on the readings presented in Figure 10, a second set of figures adapted from the schema proposed in Herbst et al. (2004) will be shown through Appendices I1-A and I1-B. As stated in Section 3.1, the valency dictionary for PDE serves as a model in the absence of one for ME. Then, the valency will be presented alongside examples retrieved from the 13th century's sampled subcorpus in List 3 (see Appendix J).

Now that the readings found for the focused-upon century's subcorpus were presented through Appendices I1-A and I1-B with examples in List 3 (Appendix J), some considerations are essential. In order to avoid exhaustiveness through repetition, the reader should consider the decisions taken for the analysis of the 12th century's sampled lexical occurrences. As pointed out for the 12th century (see Section 3.1), no occurrence of *haven* in the passive voice was attested in the 13th century as well. By the way, passive voice was quite rare in Middle English[32]. In addition, due to the French influence in the English language after the Norman Conquest, double negation can be observed, for instance, in the only occurrence whose diathesis is "[N]" coded as *1b(a) – 'to have (s. or sth. somewhere)'*. In this sentence, *haven* is in the negative form, and even so the negative particle *non* is present, as shown below.

[32] In fact, according to Toyota (in Díaz Vera, ed., 2002, p. 574), passive construction in Middle English took place periphrastically with the verbs *be* ("be") and *worthe* ("be, become"), the former being the commonest and the latter becoming old-fashioned in the 14th century. The exception is *hoten*, ("call"), morphological passive *hat(t)e*. No passive with *haven*, morphologically or periphrastically, is attested. Additionally, Lamont (2005) shows that passive voice is an innovation in the English language by saying "[p]art of the reason why the progressive passive was so late in developing is related to the development of the English passive in general. **Old English had no syntactic or morphological passive.** As a result, the sentence "The man killed" in Old English could potential mean the active "the man killed (someone)", or the passive "the man (was) killed". Context usually informed the reader. ME often relied on similar extrapolations, which Visser [1970] calls "passival" and others call "middle voice": passive in meaning, but lacking explicit passive form" (Item 5, available at: http://homes.chass.utoronto.ca/~cpercy/courses/6362-lamont.htm).

1b(a) – 'to have (s. or sth. somewhere)'
[N] – *suche wede that naveth king ne kayser non*
such wedding that had_not king not emperor not
"[He brought you to] such wedding that had neither king nor emperor"

Additionally, relative clauses were also found, in which the relative particle is duplicated (e.g. Arg1 followed by Arg1').

List 3 (Appendix J), introducing one example for each diathesis found for each reading, shows that: 1) *haven*'s diathesis "[N]" requires a Subject (Arg1) and a Direct Object (Arg2); 2) "[N PP]" requires a Subject (Arg1), a Direct Object (Arg2) and an Indirect or an Oblique Object (Arg3); 3) Infinitive clauses can consist of one of *haven*'s arguments, taking place alongside verbal particles such as 'to' or without it, as well as, in some cases, be inferred by the context – just as stated for the analysed occurrences of the 12th century's sample. In respect to the morphology of infinitive *haven,* forms without the typical infinitive ending *–en* are attested (e.g. *habbe*); 4) Clauses introduced by conjunctions such as "as" and "that" (and their spelling variants) also took place as one of *haven*'s arguments.

Valency presented and considerations made, it is now time to introduce the data retrieved from S01 and S02, and then go back to S03. As for the former, Table 9 will show the most common forms of *haven* found in the 13th century's subcorpus as a whole (or unsampled).

Table 9. S01 for the 13th century: Most frequent forms of *haven*

TMN	Gloss	Forms of *haven*	#Occurrences	Text distribution
H0	haven-INF	have	9	B002 [2]; B007 [5]; B009 [2]
		habbe	6	B067
H1	haven-PRS.1SG.IND	have	14	B007 [3]; B010 [5]; B014 [3]; B016 [3]
		habbe	13	B030 [1]; B067 [12]
	haven-PRS.2SG.IND	hauest	13	B030 [3]; B067 [10]
		hast	4	B002 [1]; B009 [1]; B014 [1]
	haven-PRS.3SG.IND	haueþ	8	B067
		haveth	5	B004 [4];B007 [1]
		hath	4	B002 [2]; B010 [2]
	haven-PRS.1PL.IND	hauen	1	B050
		haue	1	B050
		habbeþ	1	B067
	haven-PRS.3PL.IND	habbeþ	3	B067
H1’	NEG-haven-PRS.2SG.IND	neuestu	1	B067
		nauestu	1	B067
	NEG-haven-PRS.3SG.IND	naueþ	5	B067
	NEG-haven-PRS.3PL.IND	nabbeþ	3	B067
H2	haven-PST.1SG.IND	hadde	1 occ. (only)	B067
	haven-PST.2SG.IND	haddest	1	B021
		hefedest	1	B030
	haven-PST.3SG.IND	hadde	16	B007 [1]; B009 [1]; B067 [14]
		havede	7	B007
	haven-PST.3PL.IND	hadde	2	B020
H2’	NEG-haven-PST.2SG.IND	naddest	1 occ. (only)	B067
	NEG-haven-PST.3SG.IND	nadde	2	B067
H3	haven-PRS.1SG.SBJV	habbe	3	B030
	haven-PRS.2SG.SBJV	hauest	2	B067
	haven-PRS.3SG.SBJV	hath	1	B020
		haueþ	1	B067
	haven-PRS.3PL.SBJV	habbeþ	1 occ. (only)	B067
H3’	NEG-haven-PRS.2SG.SBJV	nauest	1 occ. (only)	B030
H4	haven-PST.1SG.SBJV	hadde	1	B010
		heuede	1	B030
	haven-PST.2SG.SBJV	heuedest	1 occ. (only)	B030
	haven-PST.3SG.SBJV	hadde	1 occ. (only)	B067
	haven-PST.3PL.SBJV	hadde	1 occ. (only)	B067
H5	haven-PRS.2SG.IMP	have	5	B007 [2]; B009 [1]; B010 [1]; B021 [1]

Table 9 illustrates the occurrences of the variant forms of *haven* by gloss and text distribution. The hapax legomena are of least importance, except when the corresponding form took place only once for a given gloss (e.g. *naddest* for H2’) or when more than one hapax legomenon overlapped for the same gloss (e.g. *hadde* and *heuede* for H4). So far, I will assume that the reader is already familiar with the TMN (Tense, Mode and Negation) codes, so there is no need to repeat what they stand for.

Before making any other considerations about the spelling variation in these data, it is necessary to comment the nature of the sample. Text B067 has 10,934 tokens, being the largest text of the sample. The reason to keep this text in the corpus is due to the scarce material available to compile and its removal would result in a sample smaller than the expected 20,000 tokens, though it might affect the representativeness of CMEL's subcorpus for the 13th century. Forms like *habbe* (H0, 9 occurrences), *haueþ* (H1, 8 occurrences) and most hapax legomena in, for instance, H1' undoubtedly happened to be found in this text. This fact points to this unbalance in the text distribution and, which will not be as noticeable taking into consideration all the centuries.

The same question proposed for the reader to think about why some variants for the same gloss are preferred over others presented in Section 3.1 for the 12th century remains. Since the 13th century's subcorpus presents a text that is much larger than the rest, it is expected that less forms marked as "distributed" would occur. So is the case – an example is the overlapping between the forms *have* and *habbe* for H1 – haven-PRS.1SG.IND, with 14 and 13 occurrences each distributed amongst all texts.

Let us give more attention to H0 and H1. For H0, both forms found present no *–en* ending typically assumed to correspond to the infinitive form. This illustrates the morphological impoverishment mentioned in Section 3.1 by means of Mossé (1968) words. Besides, the form *have* is distributed amongst the texts (9 occurrences) in opposition to *habbe*, which occurs six times only in B067. Having this in mind, the author's familiarity with a given spelling rather than with others for the same gloss plays a very important role here. The unknown author of B067 also employed the negative overlapping hapax legomena *nauestu* and *neuestu* for H1' – NEG-haven-PRS.2SG.IND, while others did not. In fact, all the H1' and H2' forms were found in this text. As for H3', the only form for NEG-haven-PRS.2SG.SBJV, *nauest*, was found only once, in Text B030. This profile shows that 1) the

author of B067 was perhaps more conservative in applying negative forms of *haven*, which leads to 2) the negative form of the focused-upon verb was indeed becoming obsolete, what was observed for the 12th century (see Section 3.1).

Past tense was rare. Still regarding text distribution and glosses, for H2, the third person singular presented 16 distributed occurrences of *hadde*, followed by 7 occurrences of *havede* in Text B007. Plural forms in H2 were even rarer – only *hadde* was attested twice in Text B020.

For H3, H3' and H4, only hapax legomena and a few cases of non-hapax legomena (at a maximum of 3 occurrences each) took place. This shows once again the morphological impoverishment in Middle English continuing on, this time for expressing irrealis (subjunctive). Indicative mode is more common than subjunctive and imperative, the latter represented only by *have* in singular 5 times in different texts.

As brought forward in the beginning of this Section, a hypothesis will be presented below after showing Table 10.

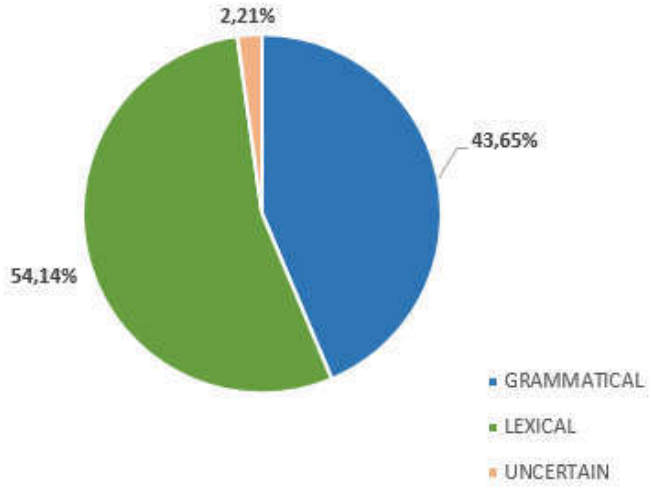
Table 10. S02 for the 13th century: Classification of the forms of *haven* (unsampled) and their respective absolute and relative frequencies

GRAMMATICAL		**LEXICAL**		**UNCERTAIN**		**TOTAL**	
Abs.	Rel.	Abs.	Rel.	Abs.	Rel.	Abs.	Rel.
79	43.65%	98	54.14%	4	2.21%	181	100%

Table 10 stands for the S02 of the 13th century. This spreadsheet proposed a classification of the forms of *haven* into three categories, namely grammatical, lexical and uncertain. Both absolute and relative frequencies are given in this table and in Graph 3.

For a further representation of the data in Table 10, Graph 4 is given below.

Graph 4. S02 for the 13th century: Relative frequencies according to their classification



As a matter of fact, if the reader compares both Table 7 (see Section 3.1 for more on the 12th century) and Table 10, or Graphs 2 and 4 for the respective centuries, one will notice an alarming fact. As the 12th century is the transitional period between Old and Middle English and a grammaticalisation process took place quite strongly even before the French influence brought to English with the Norman Conquest, one would well assume that grammaticalised *haven* forms would be more common in the 13th century's written production than in the one representing the 12th century. In other words, the more logical expectation would be that this process continued on (and, in fact, it did, as Sections 3.3 and 3.4 will show further). However, the data shown in Table 10 and Graph 4 disclaims it entirely – 54.14% of the forms of *haven* in this very subcorpus took place as lexical items (in blue) versus 43.65% as grammatical items (in green). This can be justified by the larger size of Text B067, which represents around half of the whole subcorpus. This text contains indeed many occurrences of *haven* as full-verbs expressing, for instance, the MED subsense (here called reading) *4c(b) - 'to be endowed with (a physical quality)'*. The whole context of this text consists of a debate between two characters pointing out physical characteristics in one

another, or stating what they themselves are endowed with, such as their skills. Below follows an example of 4c(b).

4c(b) – 'to be endowed with (a physical quality)'
[N] – *ich habbe gode sene*
I have good sight
"[It is said about me that] I have a good sight"

Until now, no ordering has been done regarding the most frequent readings found in the 13th century's lexical occurrences sample. This time, let us consider List 4 (Appendix K) accounting for the most frequent forms of *haven*, thus disposed in decreasing order of occurrence. In the 13th century's sampled lexical occurrences, 13 were attested as hapax legomena, including two that are not contemplated by MED. Since the number of occurrences greater or equal to 4 was considered the cutpoint to set the boundaries between the most relevant occurrences, all the other occurrences below this number were grouped with the hapax legomena (H.L.), totalising 23 occurrences.

Table 11 and Graph 5 depict the data from S03, namely the most frequent readings for the 13th century according to MED.

Table 11. S03 for the 13th century: Most frequent readings according to MED (sampled lexical occurrences)

Reading code	Reading meaning	Abs. Freq.	Rel. Freq.
4b(a)	'to have (sth. that denotes a state or condition)'	9	16.98
2(a)	'to be provided with (a part, an organ)'	5	9.43
1a(a)	'to possess'	4	7.55
4c(b)	'to be endowed with (a physical quality)'	4	7.55
4d(a)	'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'	4	7.55
7c(a)	'to obtain (sth. abstract)'	4	7.55
H.L.	Other occurrences	23	43.39
TOTAL		**53**	**100%**

According to Table 11, the most common reading is *4b(a) – 'to have (sth. that denotes a state or condition)'*, with 9 occurrences (16.98%), followed by *2(a) – 'to be provided with (a part, an organ)'*, with 5 occurrences (9.43%). Other three take place 4 times

each, namely *1a(a) – 'to possess'*; *4c(b) – 'to be endowed with (a physical quality)'*; *4d(a) – 'to experience (a feeling, fear, pleasure, wonder, joy, etc)'*, and *7c(a) – 'to obtain (sth. abstract)'*. Each one of these three last readings reached 7.55% in respect to their relative frequencies. All these readings are exemplified as follows.

4b(a) – 'to have (sth. that denotes a state or condition)'
[N] – *Havelok mihte sei þat him ne havede grip or ern*
Havelok could say that him not had skill or honour
"[Havelok could say] that he had neither skill nor honour" - Arg1 V Arg2

[N to-INF(haven)] – *me is lof to habbe reste*
me is important to have rest
"for me it is important to have rest" – Arg1 V(to-INF-haven) Arg2

2(a) – 'to be provided with (a part, an organ)'
[N] – *þu hauest wel scharpe clawe*
you have well sharp claws
"you have well sharp claws" - Arg1 V Arg2

1a(a) – 'to possess'
[N] – *hevy hameres they han*
heavy hammers they have
"they have heavy hammers" - Arg2 Arg1 V

[N INF] – *the plates that he haven wolde*
the plates that he to_have wished
"the plates that he wished to have" - Arg2 Arg2' Arg1 V

4c(b) – 'to be endowed with (a physical quality)'
[N] – *ich habbe gode sene*
I have good sight
"[It is said about me that] I have a good sight" - Arg1 V Arg2

4d(a) – 'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'
[N] – *he havede michel shame*
he had much shame
"he had much shame" – Arg1 V Arg2

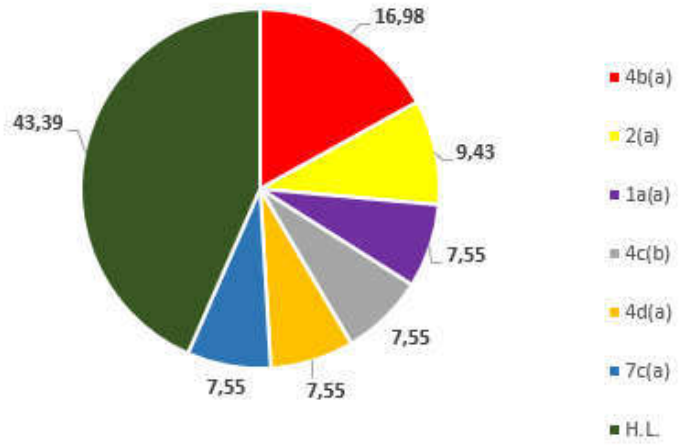
[N PP] – *luve God almighty and of him have drede*
love God Almighty and of him have fear
"love the God Almighty and have fear of Him" - Arg3 V-Arg1 Arg2

7c(a) – 'to obtain (sth. abstract)'
[N] – *ich habbe boþe luue and þonc*
I have both love and thanks
"I have both love and thanks" - Arg1 V Arg2

[N INF(haven)] – *ich bidde þat men beon iwarre an habbe gode reads*
I beg that men be aware and have good advice
"I beg men to be vigilant and have good advice" - Arg1 V(INF-haven) Arg2

As for the hapax legomena, the 23 occurrences grouped together based on the cutpoint represent 43.39% of the occurrences of the focused-upon sample. Graph 5 provides a better visualisation of the whole profile of the sample regarding the readings.

Graph 5. S03 for the 13th century: Most frequent readings according to MED (sampled lexical occurrences)



Graph 5 followed the same colour schema of Graph 3 (see Section 3.1). Dark green represents the hapax legomena group; red depicts *4b(a)*; yellow stands for *2(a)*; purple illustrates *1a(a)*; grey regards *4c(b)*; orange respects to *4d(a)*, and sky-blue shows *7c(a)*.

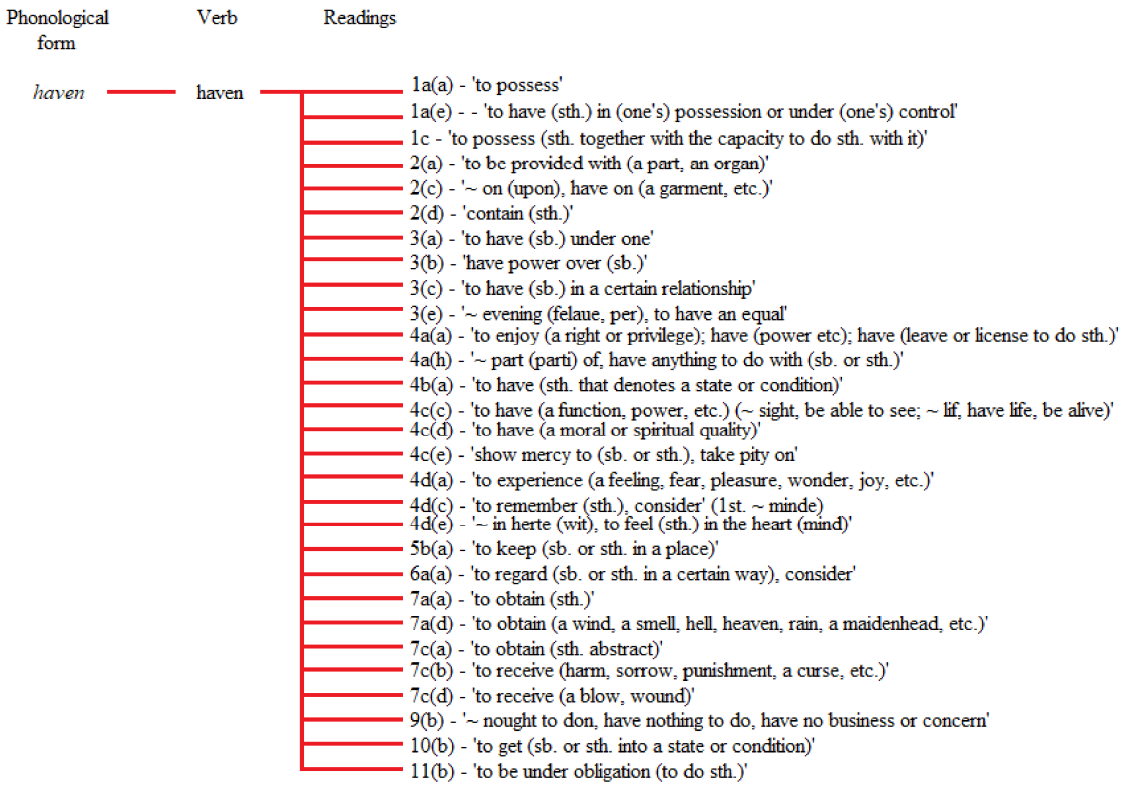
The following Section will account for the results referring to the 14th century's subcorpus.

3.3. 14th Century

As decided for Sections 3.2, 3.3 and 3.4, the present one, accounting for the 14th century's data, the skeletal structure of Section 3.1 will be followed. In the present Section, I will show and explain the results found for this very subcorpus.

The first step will be introducing the reader to the readings attested for the subcorpus of the century of interest. Figure 11 depicts the schema adapted from Perini (2015) (see Chapter 1, Figure 5).

Figure 11. Readings based in MED found in the sampled lexical occurrences for the 14th century



The purpose of introducing Figure 11 is to exhibit the readings found for the 14th century's sampled lexical occurrences. As noticeable, they are numerically ordered based on the codes of the senses and subsenses from the Middle English Dictionary (see Appendix B)

adopted here. In comparison to the ones found for 13th century (see Section 3.2), there is no undocumented sense and subsense (what I here call reading) according to MED. Then, 29 different readings were attested for the 14th century. An interesting fact about this number of readings is that, comparing it to the ones found for the previous centuries, it is higher (21 and 24 for the 12th and 13th centuries respectively). A hypothesis for this fact is perhaps that *haven*'s set of meanings and context-based usages has increased, and thus the verb *haven* gains more variety in abstraction.

Furthermore, Figure 11 leads us to Appendices L1-A and L1-B, both accounting for the adaptation of Herbst et al.'s (2004) schema in their valency dictionary for PDE. The valency found is shown in the abovementioned Appendices and each diathesis for each reading will be exemplified in List 5 (see Appendix M) afterwards.

As Section 3.1, accounting for the 12th century's results, serves as the model for the other Sections, this one is not different. Comparative considerations are essential, such as highlighting that until now no occurrence of *haven* as a full-verb in the passive voice was attested. Additionally, relative clauses continue to take place. Furthermore, it was also stated that Old and Norman French, as well as Latin indirectly, influenced the English language in many linguistic domains, such as the lexicon, morphology, syntax and semantics. Concerning the lexicon and its morphology, the impoverishment in declension and verbal inflection was already pointed out. As for syntax and semantics, especially in the genre of interest (lyrics), the flexibility in syntactic order and the nuances of senses and subsenses, as MED calls it (I can say it for *haven* itself) are evidences of linguistic change in the history of English.

Syntactic-semantic changes are of our utmost interest in this study. Being so, the diatheses found are shown in Appendices L1-A and L1-B and List 5 (Appendix M). According to these data, one can attest the following facts. [N] requires a Subject (Arg1) and a Direct Object (Arg2), exemplified by the readings *1a(a) – 'to possess'*, *2(d) – 'contain*

(sth)', *3(a) – 'to have (sb.) under one'*, *4a(a) – 'to enjoy (a right or privilege), etc'*, *4c(d) – 'to have (a moral or spiritual quality)'* and *7c(b) – 'to receive (harm, sorrow, punishment, a curse, etc.)'*.

1a(a) – 'to possess'
[N] – *thenne mot Ich have hennes arost*
then can I have hens roast
"then I can have roast hens" - Arg2 Arg1 V

2(d) – 'contain (sth.)'
[N] – *the breste had another [brod chechun]*
the chest had another [large shield-shaped_ornament]
"[With a well-fitting tunic fastened at the sides, a large shield-shaped ornament in the back]; the chest had
another" - Arg1 V Arg2

3(a) – 'to have (sb.) under one'
[N] – *and iche a segge that I see has sexe mens doke*
and each visitor that I see has six men's amount
"and each visitor I see has the amount of six men" - Arg1 Arg1' V Arg2

4c(d) – 'to have (a moral or spiritual quality)'
[N] – *she haveth sin*
she has sin
"[Unless she loves me] she has sin" - Arg1 V Arg2

7c(b) – 'to receive (harm, sorrow, punishment, a curse, etc)'
[N] – *hethyng have the hathell that any harme thynkes*
shame have the warrior who any harm thinks
"may the warrior who thinks badly of it have shame" - Arg2 V Arg1 Arg1'

Additionally, [N PP] requires a Subject (Arg1), a Direct Object (Arg2) and an Oblique or Indirect Object (Arg3). Examples are the readings *3(b) – 'have power over (sb.)'*, *3(c) – 'to have (sb.) in a certain relationship'*, *4a(h) – '~ part (parti of), have anything to do with (sb. or sth.)'*, *4c(e) – 'show mercy to (sb. or sth,), take pity on'*, *4d(c) – '~ minde, to remember (sth.), consider'*, *5b(a) – 'to keep (sb. or sth.) in a place'*, *6a(a) – 'to regard (sb. or sth. in a certain way), consider'*. They are all exemplified as follows. These diatheses are the most frequent, but others are also attested.

3(b) – 'have power over (sb.)'
[N PP] – *of maidnes meke thou hast mighte*
over maidens submissive you have power
"[Richard, source of good sense ...] you have control over submissive maidens" - Arg3 Arg1 V Arg2

4a(h) – '~ part (parti) of, have anything to do with (sb. or sth.)'
[N PP] – *late the peple and the pore hafe parte of thi silvere*
let the people and the poor have part of your silvery
"let the people and the poor ones have part of your silvery" - Arg1 V Arg2 Arg3

4c(e) – 'show mercy to (sb. or sth.), take pity on'
[N PP] – *lady ha mercy of thy man*
lady have mercy on your man
"Lady, have mercy on your man" - V-Arg1 Arg2 Arg3

4d(c) – '~ minde, to remember (sth), consider'
[N PP] – *have mynde upon my supplicacion*
have mind upon my supplication
"have my supplication in mind" - V-Arg1 Arg2 Arg3

5b(a) – 'to keep (sb. or sth. in a place)'
[N PP] – *and syþen I have in þis hous hym þat al lykez*
and since I have in this house him that all pleases
"and since I have in this house all that pleases him" - Arg1 V Arg3 Arg2

6a(a) – 'to regard (sb. or sth. in a certain way)'
[N PP] – *though men to me han onde*
though men towards me have enmity
"though men have enmity towards me" - Arg1 Arg3 V Arg2

The reading *3(e) – '~ evening (felaue, per), to have an equal'*, reproduced below, whose diathesis is an Allocutive [ALC] presents no arguments since *haven* in these sentences has existential meaning.

3(e) – '~ evening (felaue, per), to have an equal'
[ALC] – *the sonne bryght that of yelownesse hadde never pere*
the sun bright that of yellowness had never equal
"the bright sun that had no equal yellowness" - ALC

Moreover, infinitive constructions may also be arguments of *haven*, or have this verb or others in the infinitive form with verbal particle 'to' or the combination 'for to' indicating purpose. An example is the reading *7c(a) – 'to obtain (sth. abstract)'* ([N to-INF(haven)]) as follows.

7c(a) – 'to obtain (sth. abstract)'
[N to-INF(haven)] – *God yeve thee grace god happes to have*
God give you grace good fortune to have
"[Sir,] may God give you grace to have good fortune" - Arg1 Arg2 V(to-INF-haven)

Besides Objects and infinitive constructions, Adjectives and Clauses are possible as well. The readings *7c(d) – 'to receive (a blow, wound)'*, whose only occurrence is a figurative construction, and *10(b) – 'to get (sb. or sth. in a state or condition)'* contain an Adjective as one of its arguments (diathesis [N ADJ]). Examples are respectively given below.

7c(d) – 'to receive (a blow, wound)'
[N ADJ] – *open thou hast thi syde*
open you have your side
"you have your side open" - ADJ Arg1 V Arg2

10(b) – 'to get (sb. or sth. into a state or condition)'
[N ADJ] – *weping hath myn wonges wet*
crying has my cheeks wet
"crying has my cheeks wet" - Arg1 V Arg2 ADJ

For Clauses, *1a(e) – 'to have (sth.) in (one's) possession or under (one's) control'* is a good example, whose diathesis is [N PP CL], as follows.

1a(e) – 'to have (sth.) in (one's) possession or under (one's) control'
[N PP CL] – *I haf hit holly in my honde þat al desyres þurȝe grace*
I have it wholly in my hand what all you_desire Your Grace
"I have it wholly under my control what you desire, Your Grace" - Arg1 V Arg2 Arg3 CL

Diatheses requiring a Clause may not necessarily require a Direct Object, as, for instance, *4d(e) – '~ in herte (wit) etc.'* ([PP CL]), given below.

4d(e) – '~ in herte (wit), to feel (sth.) in the heart (mind)'
[PP CL] – *but in her hertes I wolde they hade hou sone that Gode hem may degrade*
but in their hearts I wished they had how straightaway that God them may bring_down
"but I wish they had it in their hearts how straightaway that God may bring them down"– Arg3 Arg1 V CL

The valency and their diatheses were shown and exemplified. The next step to present the results of this study is to give attention to the data extracted from S01 and S02 for a

moment. Table 12 will illustrate S01, showing the most common spellings of the focused-upon verb found in the 14th century's subcorpus (before the sampling process and without discriminating between grammatical an lexical categories).

Table 12. S01 for the 14th century: Most frequent forms of *haven*

TMN	Gloss	Forms of *haven*	#Occurrences	Text distribution
H0	haven-INF	have	8	C014 [1]; C019 [1]; C020 [1]; C024 [1]; C033 [1]; C040 [2]; C050 [1]
		hafe	4	C024 [2]; C025 [2]
H0’	NEG-haven-INF	n'aven	1 occ. (only)	C020
H1	haven-PRS.1SG.IND	have	31	C003 [1]; C004 [3]; C005 [2]; C012 [2]; C013 [1]; C015 [3]; C016 [2]; C017 [1]; C027 [1]; C028 [1]; C040 [6]; C041 [1]; C043 [1]; C044 [1]; C046 [1]; C051 [1]; C054 [3]
		haf	7	C054
	haven-PRS.2SG.IND	hast	11	C001 [3]; C005 [2]; C015 [1]; C016 [1]; C027 [1]; C046 [2]; C051 [1]
	haven-PRS.3SG.IND	hath	36	C001 [1]; C003 [1]; C008 [1]; C009 [4]; C010 [1]; C012 [1]; C014 [4]; C015 [3]; C016 [6]; C018 [1]; C021 [4]; C034 [1]; C036 [3]; C040 [2]; C044 [1]; C047 [1]; C051 [1]
		has	6	C023 [1]; C024 [1]; C050 [4]
		hase	5	C023 [4]; C024 [1]
	haven-PRS.1PL.IND	haven	1 occ. (only)	C020
	haven-PRS.2PL.IND	have	2	C039 [1]; C042 [1]
	haven-PRS.3PL.IND	han	11	C001 [1]; C005 [1]; C006 [1]; C009 [2]; C020 [2]; C036 [1]; C051 [3]
		have	4	C027 [1]; C036 [1]; C050 [1]; C051 [1]
H1’	NEG-haven-PRS.1SG.IND	n'ave	2	C012 [1]; C041 [1]
	NEG-haven-PRS.3SG.IND	nath	3	C020 [1]; C021 [1]; C051 [1]
H2	haven-PST.1SG.IND	hade	2	C012 [1]; C020 [1]
	haven-PST.3SG.IND	had	6	C008 [1]; C023 [2]; C050 [2]; C054 [1]
		hade	5	C020 [1]; C054 [4]
	haven-PST.2PL.IND	haden	1 occ. (only)	C024
	haven-PST.3PL.IND	had	2	C025
H2’	NEG-haven-PST.2PL.IND	n’ade	1 occ. (only)	C025
H3	haven-PRS.1SG.SBJV	ha	1 occ. (only)	C015
	haven-PRS.2SG.SBJV	have	1	C012
		haste	1	C024
		haf	1	C054
	haven-PRS.3SG.SBJV	hath	6	C011 [3]; C020 [1]; C036 [1]; C051 [1]
	haven-PRS.3PL.SBJV	hath	1 occ. (only)	C024
H4	haven-PST.1SG.SBJV	hadde	1 occ. (only)	C043
	haven-PST.3SG.SBJV	hadde	2	C014 [1]; C051 [1]
		had	2	C053 [1]; C054 [1]
	haven-PST.3PL.SBJV	haden	1	C054
		hade	1	C054
H4’	NEG-haven-PST.3SG.SBJV	nade	1 occ. (only)	C053
H5	haven-PRS.2SG.IMP	have	5	C002 [1]; C003 [1]; C005 [1]; C017 [1]; C034 [1]
	haven-PRS.2PL.IMP	have	1 occ. (only)	C042
H6	haven-PRS.PTCP	havende	1 occ. (only)	C024*
H7	haven-PST.PTCP	y-had	1 occ. (only)	C018

In the 14th century’s subcorpus (unsampled), the reader must note that most forms according to their glosses are attested as distributed amongst texts, as opposed to the issue pointed out and turned into a hypothesis for the 13th century in Section 3.2. This is due to the

14th century's subcorpus nature: there is no large text representing around a half or a third of the whole subcorpus.

Infinitive forms found representing H0 have lost their morphological marking *–en*, which usually indicates infinitive. Such ending is attested only in H0', in the negative infinitive *n'aven*, which occurred only once in Text C020. As for the 13th century (see Section 3.2), the 14th century's subcorpus does not contain the typical ending *–en*, but only in the negative form. A hypothesis is that, given the fact that negative forms were rarer, the typical ending for infinitive was maintained in more obsolete forms, and, on the other hand, linguistic change such as the loss of the ending *–en* took place in the most frequent ones, namely the declarative.

Speaking of negative forms, H0' was already presented as a hapax legomenon. For H1', only first and third person singular were attested as negative forms (*n'ave* and *nath*, 2 and 3 occurrences respectively). One would expect that they would take place in a single text each, but these occurrences were distributed. H3' presented no negative form, perhaps because irrealis (subjunctive) was rarer than indicative. H4' has only a hapax legomenon in third person singular in the present subjunctive, in Text C053.

Irrealis (subjunctive) was mentioned and now deserves some attention, probably due to its rareness. In H3, the only non-hapax legomenon is third person singular in present subjunctive (represented by the gloss haven-PRS.3SG.SBJV), *hath*, with 6 distributed occurrences. Apart from that, all the other occurrences in H3 are hapax legomena. In H4, first person singular in past subjunctive consists of the hapax legomenon *hadde*, in Text C043. For third person singular and third person plural, two forms overlap for each. The difference is that the former has two forms, *hadde* and *had*, probably variants of the same form, each occurring twice and distributed in two texts, whereas the latter has two overlapping hapax

legomena in the same text (C054), *haden* and *hade*. Therefore, irrealis (subjunctive) is indeed rare, as pointed out for the other centuries in the previous Sections.

Other mode attested is imperative. For both second person singular and plural, the same form is employed, namely *have*. In second person singular, it occurs 5 times in distributed texts, as opposed to the plural form, which consists of a hapax legomenon in Text C042. This shows that the narrator addresses a single entity instead of more than one simultaneously. Usually, in CMEL's 14th century subcorpus, the nature of the texts is romantic or devotional, so this entity may be a deity such as God, Jesus Christ or the Virgin Mary, or a beloved person, in most of the times, a maiden.

H6 and H7 stand for present and past participles respectively. They are both hapax legomena. Note that in H6 the only occurrence, *havende*, in Text C024, is marked. This is because this form was employed not as a verb, but as a noun synonym for "wealth". Participles usually receive declension instead of verbal inflection for they stand between two categories, nouns and verbs. This was quite common in highly inflected languages like Classical Latin[33]. Moreover, for H7, the hapax legomenon *y-had*, originated from Old English *gehæfd* (*y-* corresponding to *ge-*), was found only in Text C018. Given that both participle forms presented consist of hapax legomena, as well as the fact that in previous centuries (see Sections 3.1 and 3.2) they were already rare and not even found in CMEL.

So far, I have introduced the leader to the readings found in the 14th century's sampled subcorpus according to our interests, then to the diatheses with examples. Besides, based on S01, the most common forms according to their glosses and a discussion about

[33] Classical Latin had six declensions, namely nominative, accusative, dative, genitive, ablative and vocative, all inflected in masculine/feminine and neuter. The last one can be set aside for it does not change in form compared to nominative. This language had three tenses – present, perfect (past) and future. For our interest here, I will consider only the present participle. Classical Latin's nominative ending was *–ns*, corresponding to *–nd* in Old and Middle English. Therefore, for the verb *habere*, "to have", *habens* means "wealth, possession" or "having", in the case of its employment as an adjective.

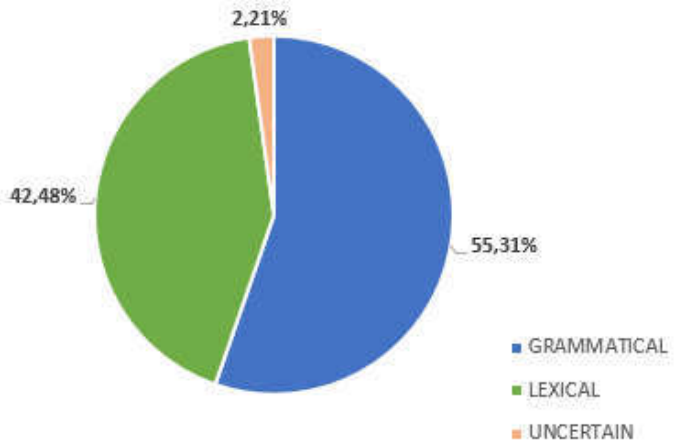
commonness or rareness of a given gloss or form were shown. What follows is the data retrieved from S02, represented by Table 13 below.

Table 13. S02 for the 14th century: Classification of the forms of haven (unsampled) and their respective absolute and relative frequencies

GRAMMATICAL		LEXICAL		UNCERTAIN		TOTAL	
Abs.	Rel.	Abs.	Rel.	Abs.	Rel.	Abs.	Rel.
125	55.31%	96	42.48%	5	2.21%	226	100%

Table 13 represents the classification of the forms of *haven* from the S01 into three categories, namely grammatical, lexical and uncertain. Reasons for this third category were clarified before (see Chapter 2 and Section 3.1). The absolute and relative frequencies for each one are given and then illustrated in Graph 6.

Graph 6. S02 for the 14th century: Relative frequencies according to their classification



Through Table 13 and Graph 6, the reader will easily notice that, apart from the uncertain, grammatical occurrences (in blue) represent more than a half of the total occurrences in the 14th century's subcorpus (55.31%). Lexical occurrences (in green) are less common, but still worthy of attention, representing 42.48% of this subcorpus' occurrences of *haven*. If the reader wishes to consider the hypothesis brought in Section 3.2 for the 13th

century, here this hypothesis is reinforced. This is so because, as stated for the 12th century about linguistic change and the process of glammaticalisation, they really tend to continue on in the language. Comparing the 12th and 14th centuries according to CMEL's data, the relative frequencies of grammatical occurrence are very close – 55.40% and 55.31% respectively. A significant difference between the relative frequencies for the lexical occurrences for each century is due to the incidence of uncertain occurrences. For the 12th century, 5.03% of the occurrences of *haven* were classified as uncertain in respect to its use versus 2.21% for the 14th century. Being so, for the 14th century, 42.48% of the occurrences was classified as lexical, and, for the 12th century, this number is lower (39.57%).

Back to the readings found in the sample extracted from the 14th century's subcorpus, they were organised based on their frequency of use, as List 6 shows (see Appendix N).

In the 14th century's sampled lexical occurrences, as Appendix N shows, 18 were attested as hapax legomena. Considering the cutpoint (greater or equal to 4) decided in order to set the boundaries between the most relevant occurrences, all the other occurrences below this number were grouped with the hapax legomena (H.L.), totalising 36 occurrences.

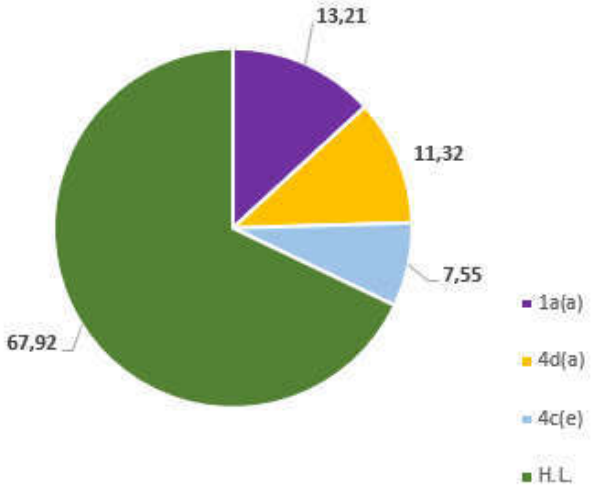
Table 14 Table 14 sums up List 6's most frequent readings, followed by their absolute and relative frequencies.

Table 14. S03 for the 14th century: Most frequent readings according to MED (sampled lexical occurrences)

Reading code	Reading meaning	Abs. Freq.	Rel. Freq.
1a(a)	'to possess'	7	13.21
4d(a)	'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'	6	11.32
4c(e)	'show mercy to (sb. or sth.), take pity on'	4	7.55
H.L.	Other occurrences	36	67.92
TOTAL		**53**	**100%**

Graph 7 below follows Table 14.

Graph 7. S03 for the 14th century: Most frequent readings according to MED (sampled lexical occurrences)



Based on List 6, Table 14 and Graph 7, the most frequent readings are, in decreasing order: *1a(a) - 'to possess'* –, in purple; *4d(a) – 'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'* – in orange, and *4c(e) – 'show mercy to (sb. or sth.), take pity on'* – in light blue. As for the hapax legomena group, in dark green, the 36 occurrences represent 67.92% of the sampled lexical occurrences of the 14th century's subcorpus of CMEL. In comparison to the readings found for the previous centuries, for the 14th century only three are more common considering the cutpoint, in opposition to five for the 12th century and six for the 13th century. As the 13th century presented the highest number of readings so far, this could be due to the hypothesis presented in Section 3.2 – one larger text can have an unusual impact in a given corpus or subcorpus.

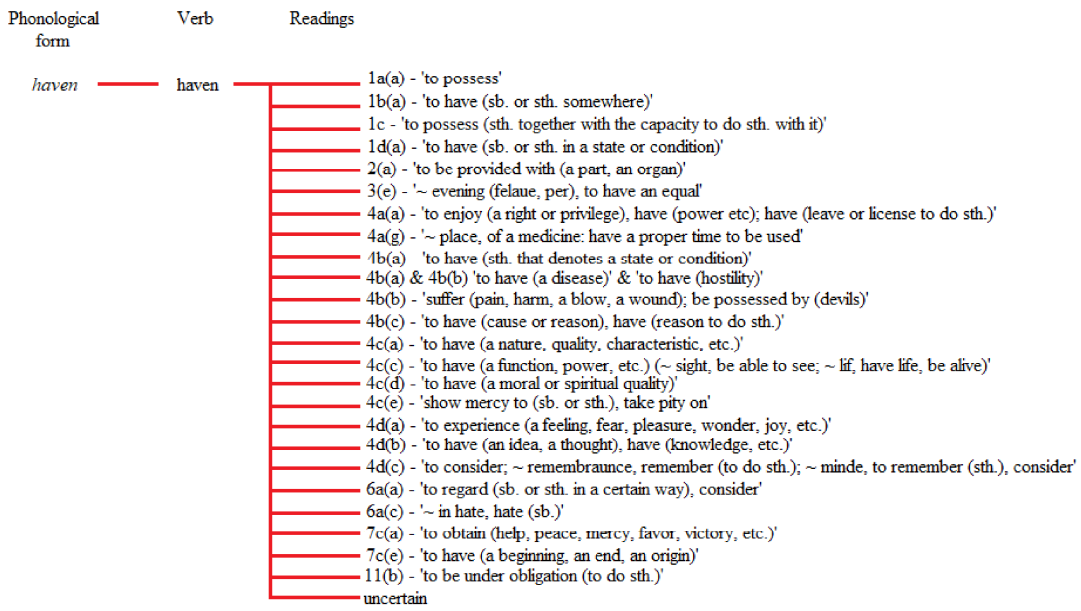
Section 3.4 will account for the results retrieved from the analysis of the data for the 15th century's subcorpus.

3.4. 15th Century

For this Section, the decisions made and explained in Section 3.1, accounting for the results for the 12th century's subcorpus, will be kept. I hereby aim to introduce the reader to the results achieved for the data retrieved from the 15th century's subcorpus of CMEL.

As for the other centuries covered by this study in the previous Sections in this Chapter, the schema adapted from Perini (2015) (see Chapter 1, Figure 5) will show, through Figure 12, the readings found for the sample of lexical occurrences of *haven* in the focused-upon subcorpus.

Figure 12. Readings based in MED found in the sampled lexical occurrences for the 15th century



Precisely as done for in the previous Sections, this schema stands for the readings found for the 15th century's sampled lexical occurrences. Naturally, they were organised in numerical order following the codes for each sense and subsense of *haven* according to MED (see Appendix B). In contrast to the readings found for the 13th century's sample, there is no

undocumented sense and subsense (as known as reading here) based on MED which I named myself. However, an interesting reading was marked as “uncertain” and there is a good reason for that. This is a particular occurrence of *haven* in the passive voice, with, as its name suggests, an uncertain classification into senses and subsenses covered by MED. As a hypothesis, one must consider the previous statements about the rareness of passive voice in Middle English. As the 15th century marks the transitional period between late Middle English and Early Modern English (EME), passive voice might have been an innovation in the language, flourished in this particular transition and originating the first attestations of the passive voice formation that continued on until Present-Day English.

In total, 25 readings were attested as noticeable in Figure 12. Section 3.3, accounting for the 14th century, showed 29 readings, and in that Section I opposed this number to the ones for the other centuries – 21 for the 12th century, and 24, for the 13th century. Then, I discussed the hypothesis of more senses and subsenses/readings to have been arisen in context-based situations and the language to have gained more variety in meaning. In the case of the 15th century, one must consider the nature of this subcorpus. As CMEL consists of a convenience sample, as stated in Chapter 2, for the 15th century more religious texts were found for the compilation. This fact may have influenced the higher incidence of some readings, as I will present further.

For now, the attention will be driven to the schema adapted from Herbst et al. (2004). Appendices O1-A and O1-B stand for the readings found and shown in Figure 12 with their respective diatheses. Then, List 7 will account for the examples for each one (see Appendix P).

As I have been doing so far, once again comparisons are necessary. In the beginning of this Section, I brought forward a hypothesis about the occurrence in the passive voice with the verb *haven*, namely the one whose reading was classified as “uncertain”. In the previous

Sections, no occurrences of this verb (or any other) was attested in the passive voice. Given the rareness of this construction, it deserves some special attention. In the last occurrence in List 7 (Appendix P), this verb, inflected in past tense, requires a Subject (Arg1) and a participle functioning as an Adjective (ADJ). The construction 'had to' could be translated with the PDE verb 'must' indicating an obligation – in this context, something extremely necessary to be done. This diathesis seems "fuzzy" precisely due to the rareness of this incidence and needs a deeper clarification in further studies. For now, it suits best to leave this occurrence as it is. Below the reader can have a look at this occurrence.

UNCERTAIN (in the passive voice)
[ADJ] – *to save mankynd ellis al the world had be forlore*
to save mankind else all the world had be forfeit
"to save mankind all the world had to be forfeit" - Arg1 V(haven) ADJ

Apart from that, the other diatheses shown in Appendices O1-A and O1-B, based on the schema in Figure 12 and exemplified in List 7 (Appendix P), will be of our most concern. Being so, most assumptions made for the other centuries seem similar, such as the fact that the diathesis [N] requires a Subject (Arg1) and a Direct Object (Arg2). For instance, in the readings *1a(a) – 'to possess'*, *2(a) – 'to be provided with (a part, an organ)'*, *4b(a) – 'to have (sth. that denotes a state or condition)'*, *4d(b) – 'to have (an idea, a thought), have (knowledge, etc.)'*, *7c(e) – 'to have (a beginning, an end, an origin)'*, among others. The examples follow.

1a(a) – 'to possess'
[N] – *y am lafte here as a woman forsake þat no goode has*
I am left here as a woman forsaken who no wealth has
"I am left here as a forsaken woman that has no wealth" - Arg1 Arg1' Arg2 V

2(a) – 'to be provided with (a part, an organ)'
[N] – *and also erys þu hast to here*
and also ears you have to hear
"and also ears you have to hear" - Arg2 Arg1 V

4b(a) – 'to have (sth. that denotes a state or condition)'
[N] – *in þat hyþe place þu hast hovnowre*
in that heavenly place you have honour
"in that heavenly place you have honour" - Arg1 V Arg2

4d(b) – 'to have (an idea, a thought), have (knowledge, etc.)'
[N] – *if þu be a lytill chyld ȝitt may þu haue þi wyll*
if you are a little child yet may you have your consciousness
"if you are a little child you may yet have your (own) consciousness" - Arg1 V Arg2

7c(e) – 'to have (a beginning, an end, an origin)'
[N] – *vnto þat ioye whyche nevyr schal have ende*
unto that joy which never shall have end
"unto that joy which shall never have an end" - Arg1 Arg1' V Arg2

For [N PP], besides Arg1 and Arg2, a third argument, an Oblique or Indirect Object (Arg3) takes place. Examples are the readings *1b(a) – 'to have (sb. or sth. somewhere)'*, *1d(a) – 'to have (sb. or sth. in a state or condition)'*, *6a(c) – '~ in hate, hate (sb.)'*, *7c(a) – 'to obtain (sth. abstract)'*, among others, exemplified below.

1b(a) – 'to have (sb. or sth. somewhere)'
[N PP] – *when ȝe play and hase your childur on kness daunsand*
when you play and have your children on knees bouncing
"[Behold, women,] when you play and have your children bouncing on your knees" - Arg1 V Arg2 Arg3

1d(a) – 'to have (sb. or sth. in a state or condition)'
[N PP] – *whanne age haþ us at his auauntage*
when age has us at its advantage
"then we may not do a big deal, [but sometimes groan, and sometimes complain, and sometimes scratch itching pustules] when age has us at its advantage" - Arg1 V Arg2 Arg3

6a(c) – '~ in hate, hate (sb.)'
[N PP] – *oure frendys þat schul loue vs beste þan wol haue vs but in hate*
our friends who shall love us dearly then want to_have us but in hate
"our friends who should love us dearly will then have us but in hate" - Arg1 Arg1' V(INF(haven) Arg2 Arg3

7c(a) – 'to obtain (sth. abstract)'
[N PP] – *thou woful moder and mayde hadest deþe in þy dolour*
you afflicted mother and maiden you_had death in your suffering
"o you, mother and maiden full of grief, who had death in your suffering" - V-Arg1 Arg2 Arg3

Infinitive clauses with or without the verbal particle 'to' (no 'for to' was attested in this sample) are attested as potential arguments (to-INF and INF) and the verb in infinitive

may be *haven* or any other. To exemplify, I cite and show the readings *1c – 'to possess (sth. together with the capacity to do sth. with it)'*, *4a(a) – 'to enjoy (a right or privilege), have (the right to possess sth.); have (power, etc.); have (leave or license to do sth.)'*, *4a(g) – '~ place; of a medicine: have a proper time to be used'*, *4b(c) – 'to have (cause or reason); have (reason to do sth.)'*, *11(b) – 'to be under obligation (to do sth.)'*, etc.

1c – 'to possess (sth. together with the capacity to do sth. with it)'
[N to-INF] – *þer he hath pour to reyse and rowte*
there it has power to raise and bellow
"it has power to raise and bellow" - Arg1 V Arg2 to-INF

4a(a) – 'to enjoy (a right or privilege), have (the right to possess sth.); have (power, etc.); have (leave or license to do sth.)'
[N to-INF] – *shall we neuer haue lycence to lyve yn ese*
shall we never have licence to live in peace
"we shall never have licence to live in peace" - Arg1 V Arg2 to-INF

4a(g) – '~ place; of a medicine: have a proper time to be used'
[N to-INF] – *I haue no place to represse þem aright*
I have no time to repress them properly
"I have no time to repress them [the wounds] properly" - Arg1 V Arg2 to-INF

4b(c) – 'to have (cause or reason); have (reason to do sth.)'
[N to-INF] – *what cause hast þu so sore to wepe*
what cause have you so sadly to cry
"what reason do you have to cry so sadly?" - Arg2 V Arg1 to-INF

11(b) – 'to be under obligation (to do sth.)'
[to-INF] – *lecchery clennesse hat mad to fle*
lust cleanness had madly to flee
"[... mercy for mankind is put away,] cleanness of lust had to flee madly" - Arg1 V to-INF

Moreover, clauses are also possible as arguments of this predicate. The only occurrence is represented by the reading *4d(a) – 'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'* as follows.

4d(a) – 'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'
[N CL] – *a secunde ioy hadyst þu whan cryst ihu of þe was borne*
a second joy had you when Jesus Christ of you was born
"a second joy you had when Jesus Christ was born of you" - Arg2 V Arg1 CL

Relative clauses remain possible as well. Besides, depending on the reading, collocations can be more or less common. The readings *3(e) – '~ evening (felaue, per), to have an equal'* and *4c(e) – 'show mercy to (sb. or sth.), take pity on'* serve as very good examples. This can be observed for the other centuries as well, irrespective of the frequency of these readings in each subcorpus. Regarding 3(e), as pointed out for the 14th century's sample results (see Section 3.3), an Allocutive (ALC) takes place but the verb itself presents no arguments. In the focused-upon century, its diathesis is empty (marked with []) having only a Subject (Arg1) as shown below.

3(e) – '~ evening (felaue, per), to have an equal'
[] – *if ȝe luf youres [son] myne has no pere*
if you love your [son] mine has no equal
"if you love your son, (then) mine has no equal" - Arg1 V

Now that the valency and their respective diatheses were shown and exemplified, I must go further and present the data extracted from the spreadsheets S01 and S02 and consider the whole 15th century's subcorpus instead of focusing on the sampled lexical occurrences.

Table 15 will account for S01, in other words, taking spelling variation into consideration and pointing out the most common forms of *haven* according to their glosses.

Table 15. S01 for the 15th century: Most frequent forms of *haven*.

TMN	Gloss	Forms of *haven*	#Occurrence	Text distribution
H0	haven-INF	have	33	D003 [1]; D006 [4]; D007 [1]; D010 [2]; D011 [1]; D017 [1]; D020 [1]; D022 [1]; D024 [1]; D025 [1]; D031 [4]; D032 [4]; D033 [1]; D034 [2]; D035 [3]; D036 [2]; D037 [1]; D038 [2]
H1	haven-PRS.1SG.IND	haue	13	D007 [1]; D008 [1]; D025 [4]; D031 [2]; D034 [2]; D035 [2]; D036 [1]
	haven-PRS.2SG.IND	hast	8	D005 [1]; D007 [1]; D016 [1]; D019 [1]; D025 [1]; D034 [3]
	haven-PRS.3SG.IND	hath	14	D002 [1]; D006 [1]; D017 [1]; D020 [2]; D025 [1]; D031 [5]; D036 [1]; D037 [2]
		has	11	D004 [1]; D006 [2]; D007 [2]; D012 [1]; D021 [3]; D035 [2]
	haven-PRS.1PL.IND	haue	2	D032 [1]; D038 [1]
	haven-PRS.2PL.IND	hase	1	D007
		haue	1	D010
	haven-PRS.3PL.IND	haue	5	D020 [1]; D023 [1]; D025 [1]; D031 [2]
H2	haven-PST.1SG.IND	hade	3	D030
	haven-PST.2SG.IND	hadest	2	D011
	haven-PST.3SG.IND	had	9	D010 [1]; D012 [3]; D024 [1]; D029 [1]; D033 [2]; D034 [1]
	haven-PST.2PL.IND	had	2	D010
	haven-PST.3PL.IND	had	3	D029 [1]; D030 [1]; D033 [1]
H2’	NEG-haven-PST.3SG.IND	nad	1 occ. (only)	D024
H3	haven-PRS.1SG.SBJV	haue	2	D031 [1]; D025 [1]
	haven-PRS.2SG.SBJV	hes	2	D033
	haven-PRS.3SG.SBJV	haue	2	D003 [1]; D031 [1]
		haþ	2	D038
	haven-PRS.3PL.SBJV	han	1 occ. (only)	D038
H4	haven-PST.1SG.SBJV	had	4	D031 [2]; D035 [1]; D038 [1]
	haven-PST.2SG.SBJV	hadde	1 occ. (only)	D034
	haven-PST.3SG.SBJV	had	3	D007 [1]; D034 [1]; D035 [1]
	haven-PST.2PL.SBJV	hade	1 occ. (only)	D036
	haven-PST.3PL.SBJV	had	2	D006
H5	haven-PRS.2SG.IMP	haue	10	D006 [1]; D008 [1]; D013 [1]; D026 [1]; D027 [4]; D033 [1]; D037 [1]
H7	haven-PST.PTCP	had	5	D025 [1]; D031 [1]; D032 [1]; D033 [1]; D035 [1]

I assume the reader is already familiar with the Leipzig Glossing Rules (2015) and the codes of Time, Mode and Negation (TMN) in Table 15. Being so, what Table 15 highlights are the different spellings of the focused-upon verb according to their glosses, their number of occurrences and their distribution among the texts. As for the similar Tables created for the S01s representing the other centuries in the previous Sections, Table 15 brings on significant points. For H0, the most frequent infinitive form *have* was attested 33 times in various texts

and without the infinitive ending *–en*. There is no negative infinitive, as it was found, for instance, in the 14th century with the abovementioned ending.

Among H1, one must pay attention to morphological differences between singular and plural. On the one hand, singular forms are quite frequent (first person *haue,* 13 occurrences; second person *hath*, 8 occurrences, and the overlapping forms *hath* and *has*, 14 and 11 occurrences respectively for the third person[34]). On the other hand, plural forms are mostly *haue*, except for the hapax legomenon *hase* in Text D007 for second person, whose gloss overlaps with *haue* in Text D010. This shows that plural forms were most susceptible to become uniform rather than singular forms. In addition, no negative form was found for H1.

For H2, comprehending past indicative, two forms are noticeable – first person singular *hade* thrice in Text D030, and second person singular *hadest*, twice in Text D011 (once more, the authors of each of these texts show preference and familiarity with these spellings respectively rather than the alternative ones for each). As for the rest of H2, *had* is preferable, especially in third person singular (9 occurrences). The incidence of hapax legomena is higher in past tense than in present tense. Besides, there is one occurrence representing H2', namely the form *nad* in Text D024.

H3 and H4 stand for irrealis (subjunctive), present and past tense respectively. H3 has each form occurring twice, in one or two texts, and there is no pattern (or, due to this subcorpus's size and nature, it is not easy to assume this lack of pattern). In any case, as discussed for the previous centuries, the process of morphological loss in the subjunctive mode was in progress, and one of the reasons for that is the employment of conjunctions such

[34] According to the Middle English Dictionary used in this study, third person singular ending in *–s* or *–th*/*-þ* could vary due to diatopic reasons. As MED indicates, the Northumbrian dialect (in the north) probably favoured *–s*, whereas West Saxon (in the south) preferred *–th*/*-þ* (e.g. *hæfes* versus *hæfþ* respectively). Probably the change from the dental fricative to the alveolar fricative was firstly accepted in the north until, years later, it reached the south. This phenomenon reached Early Modern English, as in the study conducted by Nevalainen and Raumolin-Brunberg (2003, p. 140). This has to do with social class and level of literacy; West Saxon was known as the economical and political centre of England in the Middle Ages, thus the most powerful and educated region. The ending *(e)s* was preferred by non-gentry, although gentry and nobility also used it but in a less extent in the mid-16th century. This change is, thus, in sociolinguistic terms, from below.

as “if” and “when” (and their spelling variants), which could turn the sentence into an irrealis construction. In H4, one can say that *had*, *hade* and *hadde* are variants of *had*, the form which could have originated the Present-Day English *had* in past tense. In comparison to indicative forms, subjunctive ones are rarer, as for the other centuries. This is an evidence of the first steps of the linguistic change in respect to mode morphology in English. At this point, variation in indicative forms depending on person and number was not so strong as it is nowadays.

Two facts are worthy of attention as well. The first one, regarding H5, is the significant number of occurrences of the second person singular imperative form *haue* (10 occurrences). The nature of this subcorpus, compiled with religious lyrics, tells a different story in comparison to the other subcorpora for the previous centuries. Here, most texts bring the idea of asking for divine forgiveness and mercy, so sentences such as “Have mercy on me!” are quite common. The ones asked for mercy are mostly God, Jesus Christ and the Virgin Mary, religious icons, instead of more than one entity, which could have avoided the occurrence of plural imperatives. The second interesting fact is related to H7, the past participle, which lost its *y-* (originated from Old English *ge-*[35]), became *had*, precisely like the past tense (see the variants in H4). As stated in Section 3.3, participles generally receive declension instead of verbal inflection, as in Classical Latin, but the morphological impoverishment culminated in the loss of any declension at this point in the history of English. If in the previous centuries they were rare, in the transitional period between Middle and Early Modern English such form and variants would not be expected. No present participle was attested.

[35] As a matter of fact, this morphology is current in German (*Hochdeutsch*) in past participle formation, just as it was in Old English. Please compare the past participle forms (OE) *gehæfd* and (Ger.) *gehabt* and notice the similarities.

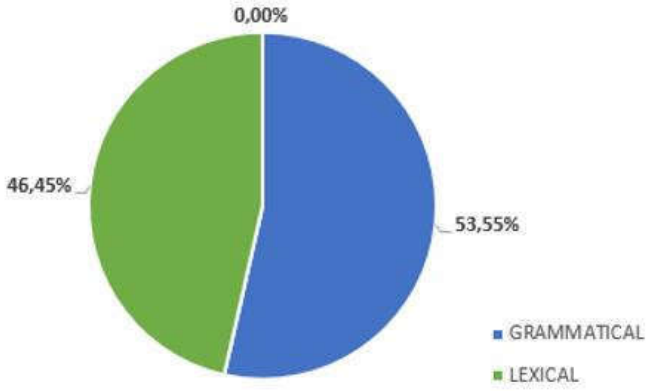
Given that the data representing the S01 for the 15th century's subcorpus was introduced and commented, it is time to go further and present the data extracted from its S02. Table 16 will account for the classification of *haven* into the three already-known categories, namely grammatical, lexical and uncertain, followed by their respective absolute and relative frequencies.

Table 16. S02 for the 15th century: Classification of the forms of *haven* (unsampled) and their respective absolute and relative frequencies

GRAMMATICAL		**LEXICAL**		**UNCERTAIN**		**TOTAL**	
Abs.	Rel.	Abs.	Rel.	Abs.	Rel.	Abs.	Rel.
98	53.55%	85	46.45%	0	0%	226	100%

In Table 16 the category called "uncertain" has no occurrence. As all the sentences and fragments containing any form of *haven* took a considerable effort to be translated with the use of MED and following the notes provided by the editors, it seems more reasonable to assume that the older the text, the higher the difficulty and effort to translate it. The texts from the 15th century are linguistically more similar to our knowledge of Present Day English and thus "easier" to comprehend and translate into PDE. More words are recognisable and present in MED, which facilitated my efforts. Moreover, in the absence of uncertain occurrences, only grammatical and lexical ones are worthy of full attention. More on this matter will be shown in Graph 8 below.

Graph 8. S02 for the 15th century: Relative frequencies according to their classification



Both Table 16 and Graph 8 show the distribution between grammatical and lexical occurrences of *haven* (in blue and light green, respectively). The relative frequency for the grammatical occurrences, 53.55%, clearly represent more than the half of the total unfiltered and unsampled occurrences, versus 46.45% of the lexical occurrences. Despite the unusual results shown for the 13th century, the process of grammaticalisation indeed was taking place severely in the history of English - which may be verified by future studies.

Returning to the filtered and sampled lexical occurrences of the 15th century's subcorpus, analysed regarding their readings based on MED's senses and subsenses of *haven*, List 8 shows below their frequency of use, organised in decreasing order (see Appendix Q).

Based on List 8 (Appendix Q), in the 15th century's sampled lexical occurrences, 12 were attested as hapax legomena, including one marked as "uncertain" for it is not contemplated by MED. Considering the cutpoint greater or equal to 4 to set the boundaries between the most relevant occurrences, all the other occurrences below this number were grouped with the hapax legomena (H.L.), totalising 35 occurrences. Table 17 represents List 8 (Appendix Q), containing the most frequent readings' absolute and relative frequencies.

Table 17. S03 for the 15th century: Most frequent readings according to MED (sampled lexical occurrences)

Reading code	Reading meaning	Abs. Freq.	Rel. Freq.
4c(e)	'show mercy to (sb. or sth.), take pity on'	10	18.87
4a(a)	'to enjoy (a right or privilege), have (the right to possess sth.); have (power etc); have (leave or license to do sth.)'	4	7.55
4d(a)	'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'	4	7.55
H.L.	Other occurrences	35	66.03
TOTAL		**53**	**100%**

Graph 9 illustrates Table 17 as follows.

Graph 9. S02 for the 15th century: Most frequent readings according to MED (sampled lexical occurrences)

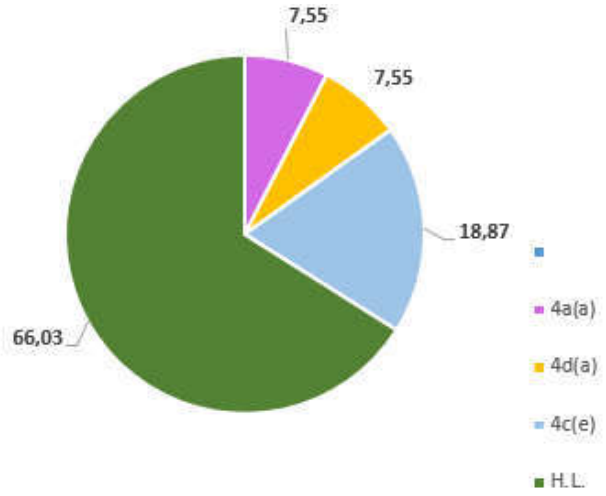


Table 17 and List 8 depict the most frequent readings, in the following decreasing order: *4c(e) – 'show mercy to (sb. or sth.), take pity on'* (in light blue), representing 18.87% of the sampled lexical occurrences. It is followed by both *4a(a) – 'to enjoy (a right or privilege), have (power, etc.); have (leave or license to do sth.)'* (in light purple) and *4d(a) – 'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'* (in orange) – each one representing 7.55% of the sampled lexical occurrences. The majority of the corpus is represented by the hapax legomena group (66.03%). In comparison to the other centuries, the

15th century presented the highest relative frequency of hapax legomena, showing a narrowing of senses and subsenses/readings in this transitional period between ME and EME.

Section 3.5 will account for further considerations on the results found for each century and a deeper comparison between them, this time having its focus on verb valency and the diatheses found for each reading for each century presented in the previous Sections.

3.5. Comparing the centuries – 12th to 15th: a focus on the readings

Chapter 1 served as a literature review on verb valency in linguistics drawing mainly on Perini (2015). His schema proved significant for the present study, which accounts for a past synchrony of the English language. Irrespective of the language and period, this schema was of great usefulness to be adapted from the senses and subsenses found in the chosen Middle English Dictionary.

As the methodological steps for the compilation of CMEL – *Corpus of Middle English Lyrics* – was exhaustively explained and discussed in Chapter 2, in this Section more attention must be given to this very Chapter. Firstly, Section 3.1 showed the results of the analysis of the 12th century's subcorpus. Secondly, Section 3.2 accounted for the 13th century's subcorpus. Thirdly, Section 3.3 dealt with the 14th century's subcorpus. Finally, Section 3.4 presented the results for the 15th century's subcorpus. A few comparisons were made over the Sections, and now they deserve in-depth comments especially in the readings and their valency. However, other data are also worthy of attention, yet they will not be the focus here.

Chapter 2 presented three types of spreadsheets, named S01, S02 and S03. Up to this point I will assume the reader is already familiar with what they account for. From now on, their data will be discussed in this order.

For the time being, let us address the first two spreadsheets. For S01, the main conclusion concerns changes in the domain of morphology. For all centuries covered by this study, I mentioned that if a given spelling variant is more common than others for the same gloss (following the Leipzig Glossing Rules, 2015), perhaps the unknown/supposed author (except for Geoffrey Chaucer, Arnold of Brescia and a few known others) was more familiar with that particular spelling and for this and other reasons he chose to employ that form. For instance, the infinitive form, represented by the code H0, 12th century's subcorpus showed *habbenn* as the most frequent, all of them found in Text A030; 13th, 14th and 15th century's subcorpora presented *have* in a variety of texts. The loss of the infinitive ending *–en* took place, and, if the reader takes a deeper look on the sentences, verbs such as *mouen* ("may", employed with the meaning of "to be able to do something, to prevail") and *shulen* ("shall", already used as a modal auxiliary) were employed and, in turn, turned the infinitive *haven* into *have*, causing the loss of its usual infinitive ending. However, cases without the abovementioned verbs, which in PDE are considered modal verbs, are also attested. This similarity is present between the 13th to the 15th centuries. A hypothesis which may sound obvious for Old English scholars is that the 12th century represents the transitional period between this synchrony and early Middle English.

Continuing the flow of this discussion, negative and irrealis (subjunctive) forms were becoming obsolete, employed mainly as hapax legomena or not even attested. Imperative forms, precisely second person singular, was more frequent in the 15th century due to the nature of this subcorpus – a compilation of available religious texts with prayers from Christians or the Virgin Mary asking for mercy and forgiveness. This attests that the context plays a very important role in the whole understanding of the text. What is at stake here is the frequency of the readings (based on MED's senses and subsenses of *haven*) depending on the convenience samples representing each century studied. This will be discussed later.

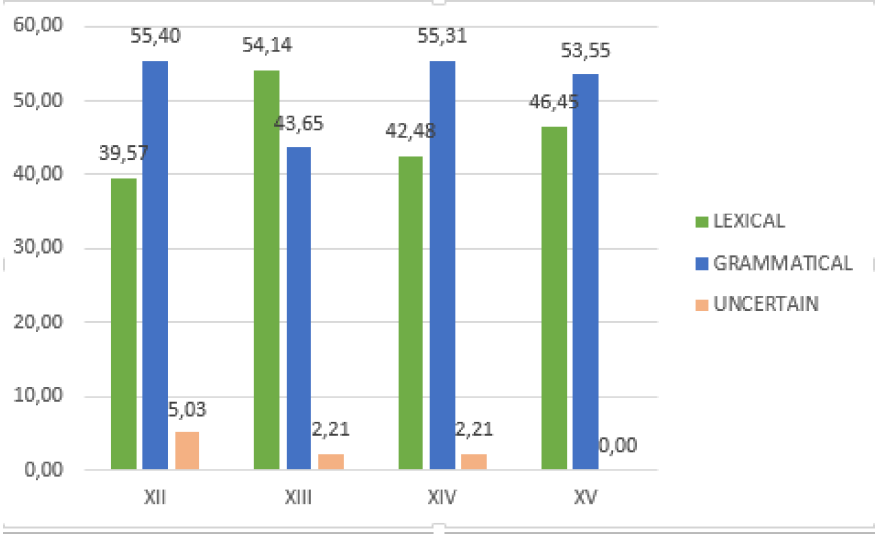
Moving to S02, I mentioned and briefly discussed the grammaticalisation process by providing references. Basically, full-verbs such as the focused-upon one can change its use, moving from the primary meaning ‘to possess, to own’ attested in Old English *habban* to others, as well as being employed as a grammatical verb to express past tense. It was thus essential to classify all the occurrences of *haven* in CMEL’s subcorpora into grammatical and lexical, and then filter and sample the latter ones. However, as I had employed much effort in the translations, some of the sentences were too obscure in both use and meaning to make possible this classification. Then, the category “uncertain” was included, meaning none of the other options. Below, I will reproduce Table 4 and Graph 1 from Chapter 2.

Table 4 - Absolute and relative frequencies of lexical, grammatical and uncertain occurrences of haven per century.

Century	**Lexical**		**Grammatical**		**Uncertain**		**Total**	
	Abs.	**Rel.**	**Abs.**	**Rel.**	**Abs.**	**Rel.**	**Abs.**	**Rel.**
XII	55	39.57%	77	**55.40%**	7	5.03%	139	100%
XIII	98	**54.14%**	79	43.65%	4	2.21%	181	100%
XIV	96	42.48%	125	**55.31%**	5	2.21%	226	100%
XV	85	46.45%	98	**53.55%**	0	0%	183	100%

Graph 1 represents the relative frequencies of Table 4.

Graph 1. Frequency of items per century.



As noticeable, the uncertain occurrences, in pink, represent the minority of all occurrences for the first three centuries – no occurrence marked as "uncertain" was attested for the 15th century's subcorpus, its reasons explicated in Section 3.4. In blue, the grammatical occurrences prevail, except for the 13th century (the hypothesis for that was already clarified in Section 3.2). For the 12th century, it can show that the grammaticalisation process, already in course since Old English, was becoming quite frequent (55.40%). It is worthy to remember that the 12th century is the transitional period between Old and Middle English. For the 14th century, this framework is attested as well (55.31%). Finally, for the transitional period between late Middle English to Early Modern English, namely 15th century, the situation is indeed similar (53.55%).

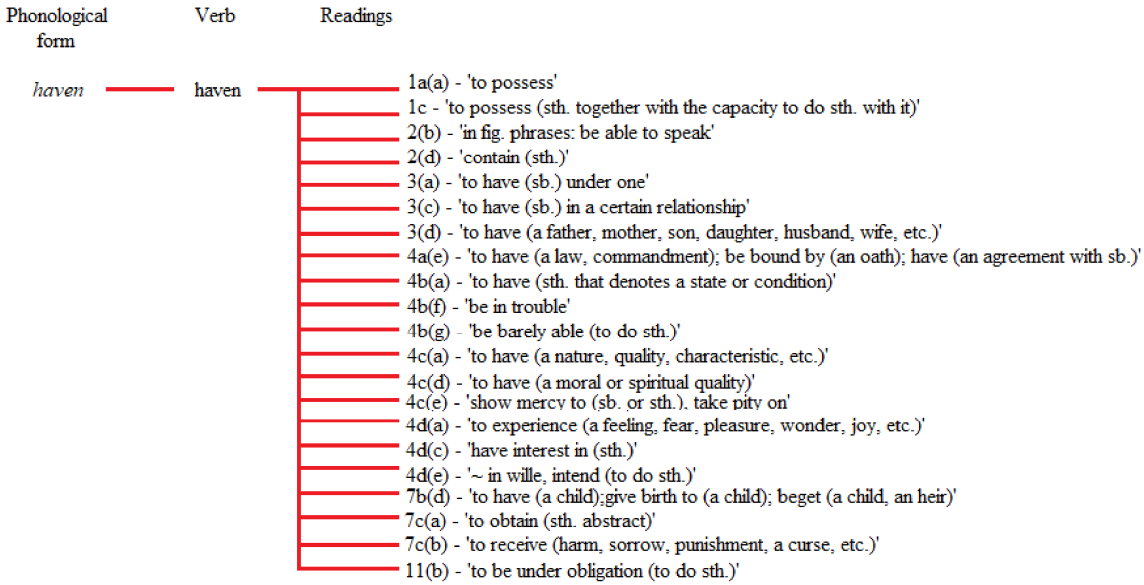
Despite the grammaticalisation process, which prevailed in English in the focused-upon verb, occurrences of *haven* classified as lexical, in light green, are considerable. For the 12th century, 39.57% represents this situation. For the 14th century, this relative frequency increases to 42.48%, and continues to grow in the 15th century, reaching 46.45%. The 13th century portraits a peculiar case due to the occurrence of a larger text representing around a

half of the whole subcorpus, so a different behaviour is naturally expected. This was a hypothesis discussed in its respective Section.

Given the well-deserved attention to the data extracted from S01 and S02 for all studied centuries, now it is worthy to focus on S03 and the readings found for each century, as well as their valency. S03 was the result of a filtering and sampling process by means of *R* software (R CORE, 2019; see Chapter 2) to sort a balanced sample of lexical occurrences to be analysed. There is no way to avoid a discussion in the domain of syntax and semantics, even because the latter influences considerably the readings, which, in turn, are context-based.

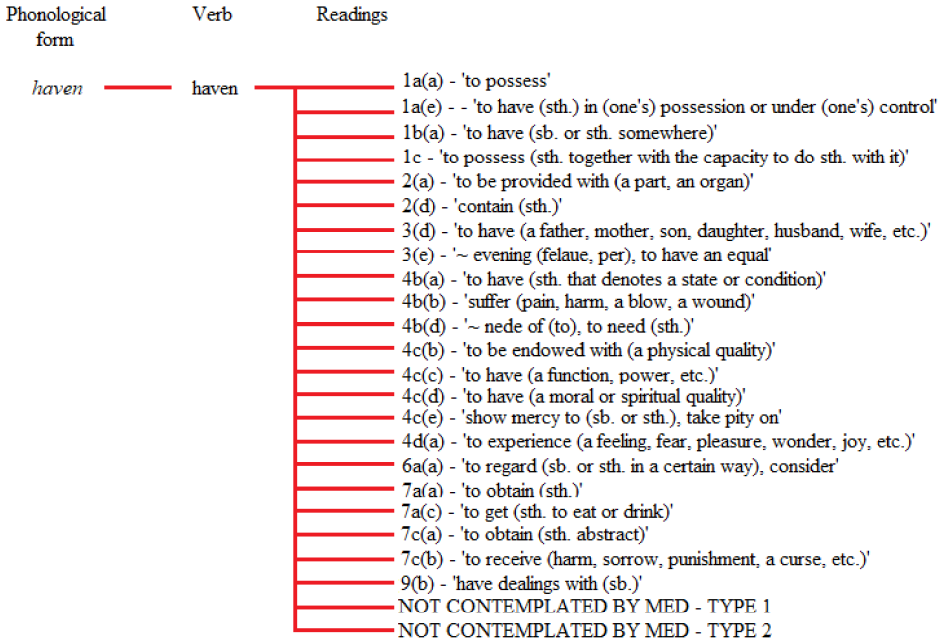
The first step in this comparison of results is to verify the readings found for each century and which ones took place in all centuries. For this purpose, I will reproduce Figures 9, 10, 11 and 12, accounting respectively for the 12th, 13th, 14th and 15th centuries in their respective Sections. Let us begin with Figure 9.

Figure 9. Readings based in MED found in the sampled lexical occurrences for the 12th century.



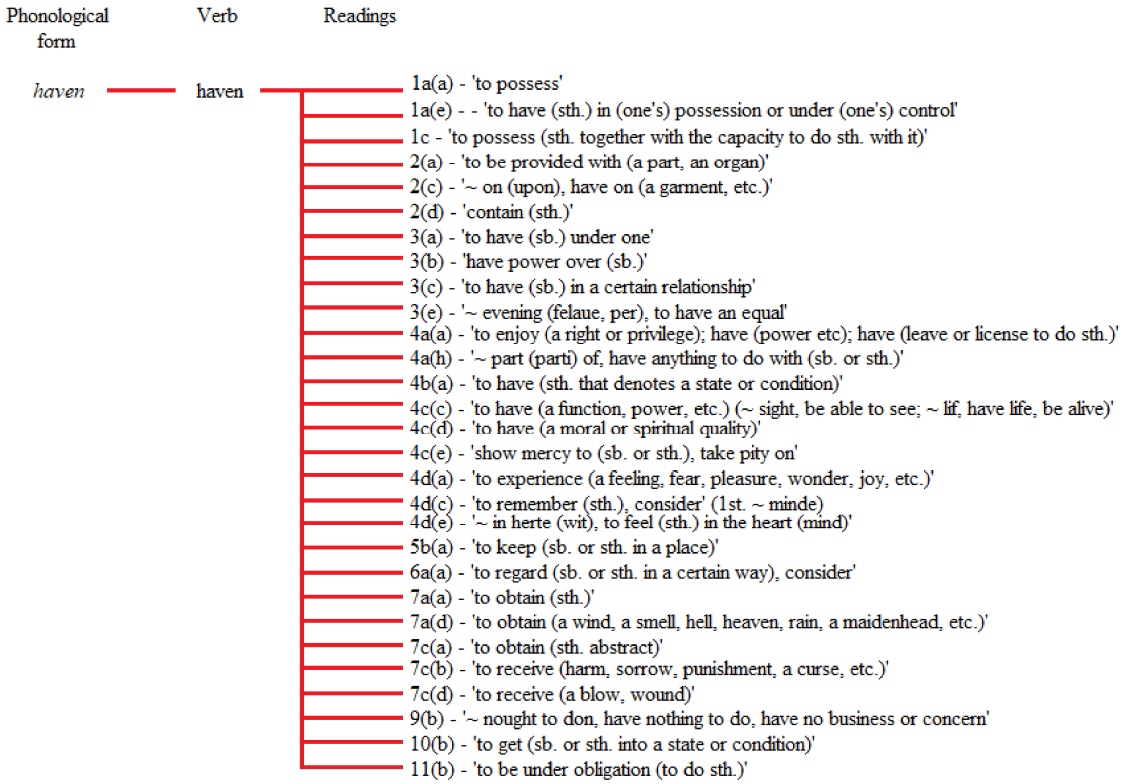
Then, Figure 10 follows.

Figure 10. Readings based in MED found in the sampled lexical occurrences for the 13th century.



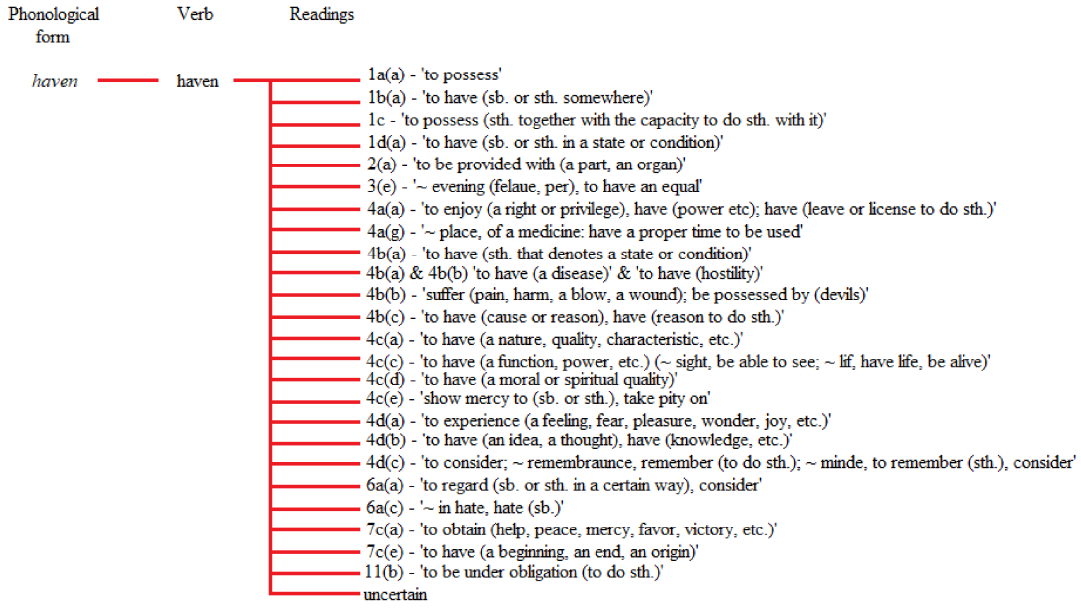
Afterwards, comes Figure 11.

Figure 11. Readings based in MED found in the sampled lexical occurrences for the 14th century.



Lastly, I present Figure 12.

Figure 12. Readings based in MED found in the sampled lexical occurrences for the 15th century.



I urge the reader to pay great attention to the Figures reproduced above. They lead us to Appendix R.

Appendix R sums up all the readings found for the sampled occurrences of *haven* in CMEL, including those not contemplated by MED and named by me and the "uncertain" one (the passive occurrence in the 15th century; for more on this, see Section 3.4). The colours are meaningful here, symbolising particular occurrences for each century or for all centuries. Light purple stands for the readings common to all centuries, namely *1a(a) – 'to possess'*, *1c – 'to possess (sth. together with the capacity to do sth. with it)'*, *4b(a) – 'to have (sth. that denotes a state or condition)'*, *4c(d) – 'to have (a moral or spiritual quality)'*, *4c(e) – 'show mercy to (sb. or sth.), take pity on'*, *4d(a) – 'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'* and *7c(a) – 'to obtain (sth. abstract)'*. In this moment, it does not matter which of them is more frequent or consists of a hapax legomenon; what is of highest importance here is their prevalence over the covered centuries.

The other colours represent the occurrences particular to a century. Light green stands for the occurrences that take place only in the 12th century's sample. Yellow, for the 13th century's. Light blue, for the 14th century's and finally light salmon for the 15th century's sample. Four readings, now considering only their respective codes, are particular of the 12th century – *4a(e)*, *4b(f)*, *4b(g)* and *7b(d)*. Regarding the 13th century, this number increases to six, as known as *4b(d)*, *4c(b)*, *5b(a)*, *7a(c)* and the two not contemplated by MED, *Type 1 ('to be quiet')* and *Type 2 ('to have an answer')*. The 14th century's sample has five particular occurrences, namely *2(c)*, *4a(h)*, *7a(d)*, *7c(d)* and *10(b)*. Lastly, eight particular occurrences were attested in the 15th century's sample, namely *1d(a)*, *4a(g)*, *4b(a) & 4b(b)*, which is peculiar due to its double context-based classification, *4b(c)*, *4d(b)*, *6a(c)*, *7c(e)* and the "uncertain" one.

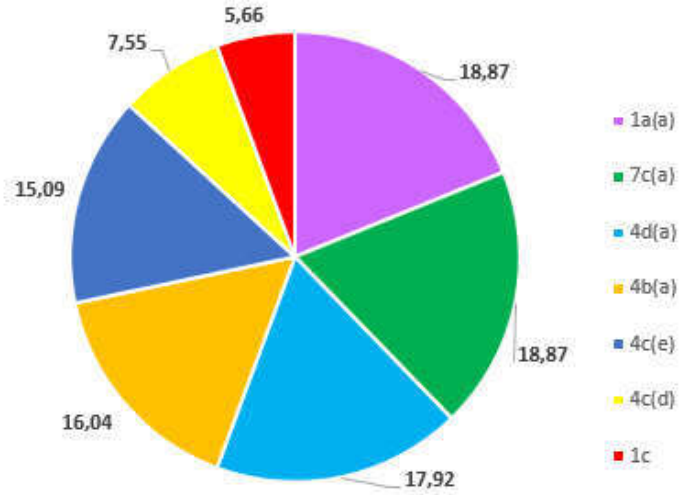
Let us now consider frequency. In each Section accounting for each subcorpus' sample, I showed in decreasing order the most frequent readings. This time, I will show through Table 18 both the absolute and relative frequencies for the common readings for all centuries as shown in Appendix R in light purple, also in decreasing order.

Table 18 – Common readings to all centuries and their absolute and relative frequencies.

Reading code	Reading meaning	Abs. Freq.	Rel. Freq.
1a(a)	'to possess'	20	18.87
7c(a)	'to obtain (sth abstract.)'	20	18.87
4d(a)	'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'	19	17.92
4b(a)	'to have (sth. That denotes a state or condition)'	17	16.04
4c€	'show mercy to (sb. Or sth.), take pity on'	16	15.09
4c(d)	'to have (a moral or spiritual quality)'	8	7.55
1c	'to possess (sth. Together with the capacity to do sth. With it)'	6	5.66
TOTAL		**106**	**100%**

Graph 10 illustrates Table 18 as follows.

Graph 10. Readings common to all centuries and their absolute and relative frequencies



In Table 18 and Graph 10 it is noticeable that the readings *1a(a)* and *7c(a)* have equal relative frequencies, followed by reading *4d(a)*, *4b(a)*, and *4cI*. Low relative frequencies are depicted for *4c(d)* and *1c*.

The first noticeable fact is that the readings *1a(a) – 'to possess'* and *7c(a) – 'to obtain (sth. Abstract)'* overlap, meaning that the primary meaning of possession attested in the etymology of *haven*, originated from Old English *habban*, is still strong at this point in the history of English. Together comes the meaning of obtaining, which leads to a causative process such as 'I have something because I obtained/I was given it'. What is interesting is that *7c(a)* brings the idea of obtaining something abstract, like help or mercy, instead of a material object. Another remarkable fact is the incidence of group 4 – *4d(a)*, *4b(a)*, *4cI* and *4c(d)*. According to MED (see Appendix B), one can infer with a keener eye that this group stands for abstractness and ephemeral or intrinsic states or conditions, such as emotions, illnesses and mental experiences. This is probably an innovation in the language, the semantic range of *haven* (or even OE *habban*) being broadened. Additionally, creative written production could perhaps enhance the chances for such uses to take place, especially considering that lyrics are meant to be sung on in the first place and metrics can be of such a great influence, otherwise the editors would not devote their time and effort to reconstruct the music behind old manuscripts.

The time to focus on the valency of *haven* has finally come. Introducing the readings found for each century's sampled lexical occurrences was of great usefulness, as I did in Sections 3.1, 3.2, 3.3 and 3.4. In the present Section, I showed through Table 18 and Graph 10 the overall most frequent readings alongside their absolute and relative frequencies. Comments were added concerning the higher incidence of group 4, as abovementioned. The matter of innovative phenomena in the English language over its history was also of great concern.

To summarise the data shown in each Section accounting for each covered century (Appendices F1-A and F1-B; Appendices I1-A and I1-B ; Appendices L1-A and L1-B , and Appendices O1-A and O1-B), the following step is to present Appendices S1, S2, S3 and S4.

With the purpose of updating the reader with the number of readings and diatheses found for each century, Appendices S1, S2, S3 and S4 are essential. The 12th century presented 11 diatheses and 21 readings, as shown in Appendix S1. Next, the 13th century had 10 diatheses and 24 readings. The hypothesis concerning the nature of the 13th century's subcorpus was presented in Section 3.2 and may have to do with the narrowing of diatheses, but the number of readings show a different framework: new context-based meanings rise. As a matter of fact, for the 14th century 14 diatheses and 29 readings were attested. Compared to the previous centuries, this one is portrayed as the most innovative in terms of more variability in the diatheses and readings. Finally, for the 15th century, a semantic narrowing process is observed – 12 diatheses and 25 readings, less than the respective numbers found for the 14th century. Again, this subcorpus has a peculiar nature. As opposed to the 14th century's subcorpus, which presented more variability of themes (e.g. love lyrics, devotional lyrics, miscellaneous lyrics and so on), the 15th century's is composed by religious lyrics. As one can notice, every subcorpus has a peculiar nature – 12th century with less variability and longer texts; 13th century with a very large text covering around a half of this subcorpus; 14th century with more variability in themes and 15th century with one particular theme. As CMEL consists of a convenience sample, this issue could not be avoided given the circumstances of compilation (as mentioned in Chapter 2).

Apart from that, the diatheses found for all four centuries in discussion must be regarded more keenly. Before starting a discussion about syntactic order of each one, separating the wheat from the chaff is crucial. In other words, the most relevant diatheses

deserve more attention, namely those that occur in all centuries, than the others particular of one or two or three centuries. In this case, the most significant ones are given in Table **19** alongside their variant syntactic orders for each century.

Table 19 – Common diatheses to all four centuries and their syntactic orders

Diathesis/Century	12		13		14		15	
	Syntax	#Occ	Syntax	#Occ	Syntax	#Occ	Syntax	#Occ
[N]	**Arg1 V Arg2**	**9**	**Arg1 V Arg2**	**12**	**Arg1 V Arg2**	**10**	**Arg1 V Arg2**	**6**
	Arg2 Arg1 V	**3**	**Arg2 Arg1 V**	**1**	**Arg2 Arg1 V**	**4**	**Arg2 Arg1 V**	**3**
	Arg1 Arg2 V	3	Arg1 Arg2 V	1	Arg2 Arg2' Arg1 V	1	Arg2 Arg2' Arg1 V	1
	Arg1 Arg1' V Arg2	**4**	**Arg1 Arg1' V Arg2**	**2**	**Arg1 Arg1' V Arg2**	**2**	**Arg1 Arg1' V Arg2**	**2**
	V Arg1 Arg2	4	Arg2 V Arg1	2	V Arg1 Arg2	1	Arg2 V Arg1	1
	Arg2 Arg2' Arg1 V	1	V-Arg1 Arg2	2	Arg1 V(INF-haven) Arg2	1		
					Arg2 V Arg1 Arg1'	1		
[N PP]	**Arg1 V Arg2 Arg3**	**3**	**Arg1 V Arg2 Arg3**	**4**	**Arg1 V Arg2 Arg3**	**1**	**Arg1 V Arg2 Arg3**	**3**
	Arg2 Arg1 Arg3 V	1	Arg1 V Arg3 Arg2	1	Arg1 V Arg3 Arg2	1	Arg2 Arg1 V Arg3	1
	Arg3 Arg1 V Arg2	1	Arg3 V Arg1 Arg2	1	Arg3 Arg1 V Arg2	1	Arg3 Arg1 V Arg2	2
	Arg1 Arg3 V Arg2	**2**	**Arg1 Arg3 V Arg2**	**1**	**Arg1 Arg3 V Arg2**	**1**	**Arg1 Arg3 V Arg2**	**1**
	Arg1 Arg1' Arg2 V Arg3	1	V-Arg1 Arg2 Arg3	1	V-Arg1 Arg2 Arg3	2	V-Arg1 Arg2 Arg3	2
			Arg3 V-Arg1 Arg2	1	Arg3 V-Arg1 Arg2	1	Arg3 V-Arg1 Arg2	1
			V Arg1 Arg2 Arg3	1	Arg1 Arg1' Arg3 Arg2	2	Arg1 Arg1' V(haven) Arg2 Arg3	1
					Arg1 Arg1' Arg3 V Arg2	1	V-Arg1 Arg3 Arg2	1
					Arg1 Arg2 V Arg3	1	Arg2 Arg3 Arg1 V	1
					Arg3 V Arg1 Arg2	1		
[N to-INF]	Arg2 Arg1 V to-INF	1	Arg2 Arg1 V Arg2 to-INF	1	Arg1 Arg1' V Arg2 to-INF	1	Arg2 V Arg1 to-INF	1
					Arg1 Arg2 V to-INF	1	Arg1 Arg2 V to-INF	4
[N to-INF(haven)]	Arg1 Arg1' V(to-INF(haven)) Arg2	1	Arg1 V(to-INF(haven)) Arg2	1	Arg1 Arg2 V(to-INF(haven))	1	Arg1 V(to-INF(haven)) Arg2	2

Notes: Arg1 stands for the Subject; Arg2, for the Direct Object; Arg3, for the Indirect or Oblique Object; to-INF for an infinitive clause with the verbal particle 'to' and, finally, to-INF(haven) for the same case, except for the fact that the verb in infinitive is *haven* itself. The colours indicate similarities.

Only the common diatheses for all four covered centuries are given in Table **19**. As they prove themselves worthy of the reader's attention, the ones whose argument structure is common for the same diathesis are highlighted in bold. The number of occurrences of each argument structure is given as well. As noticeable, no Clause (CL) or infinitive constructions with the combination 'for to' (for-to-INF), with or without *haven*, as

well as without any verbal particle, were attested as common to all centuries. They will be regarded as exceptions instead of the rule.

Speaking of rule, one might well say that, even if the genre of interest is poetry and has its own style, regularities do exist in this point in the history of English. For instance, the diathesis [N] preferred three syntactic orders over seven (the highest number, accounting for the 13th century), namely 'Arg1 V Arg2', 'Arg2 Arg1 V' and, for relative clauses, 'Arg1 Arg1' V Arg2'. For [N PP], two prevailed over ten, more specifically 'Arg1 V Arg2 Arg3' and 'Arg1 Arg3 V Arg2'. As for [N to-INF] and [N to-INF(haven)], it is hard to mention any regularity or prevalence. What can be observed is that the 14th and 15th centuries innovated with 'Arg1 Arg2 V to-INF' for the former – in the 14th century's sample, only one such occurrence was found, differently from the 15th century, which presented four occurrences. The last diathesis has few occurrences and the only resemblance is 'Arg1 V(to-INF(haven))] Arg2' for the 13th and 15th centuries, the latter presenting two occurrences whereas the former consists of a hapax legomenon.

The abovementioned framework denotes a preference of a maximum of three arguments required by the predicate *haven*. I will allow myself to return to the Chemistry-based metaphor borrowed into Linguistics by Tesnière and say that *haven* is a molecule requiring two or three slots (the former is more frequent according to Table **19**) to reach stability in the sentence – such stability means to have all necessary information for the sentence to make sense. The arguments' slots are reserved to a Subject (Arg1), Objects (Arg2 and Arg3) and infinitive clauses (to-INF, with or without *haven* inflected in infinitive). This shows that the Englishmen writing poetry, a written production which may not reproduce spontaneous speech as it is, yet is a legitimate way to represent the creative potential of the human faculties to express ideas linguistically (special attention to emotional states, such as love, fear and devotion), preferred [N] and [N PP] in their works.

As for the growing incidence for the argument structure to become more fixed, especially in [N] and [N PP], the reader should remember Mossé's (1968) words: "[d]espite the progressive impoverishment of flexion the word order in ME in which the principal elements of the sentence, subject (S), verb (V), object (O) and adjuncts (A) were placed was still very flexible […]" (p. 122). Based on that, syntactic flexibility was indeed attested (see Appendices F1-A and F1-B; Appendices I1-A and I1-B; Appendices L1-A and L1-B, and Appendices O1-A and O1-B). I call the reader's attention once again for the genre lyrics, with much more creative freedom in respect to its musical nature. However, in the common diatheses showed in

Table 19, this morphological impoverishment, the Old and Norman French influence after the Norman Conquest in the 11th century and the indirect influence of Classical Latin in English did not cause this syntactic fixity overnight. What

Table **19** shows is the very process of syntactic fixity in Middle English, first affecting [N] (Subject and Direct Object mandatory), then [N PP] (Subject, Direct Object and Indirect or Oblique Object mandatory) and then the less frequent diatheses abovementioned. The latter remained more flexible, proving Mossé's perspective. In other words, the more frequent a diathesis takes place, the quicker it is affected by linguistic change, in this case, a more fixed syntactic order.

As a more in-depth discussion focusing on the diatheses and valency variability in Middle English was driven, without excluding other details on the general results, it is the right time to move forward to the Final Remarks, in Chapter 4.

CHAPTER 4 – FINAL REMARKS

Through this Chapter, I bring the task of investigating the syntactic-semantic behaviour of the Middle English verb haven in the timespan between the 12th to the 15th century to an end. This endeavour drew on, mainly, on Corpus Linguistics. An interface between this field and Historical Linguistics was done to anchor the proposed study. Studying a past synchrony of any language with graphical registers, in other words over-centuries surviving written production, is not a quest to be completed overnight or without a strong theoretical basis and available material to compile a corpus which, itself, consists of a sample of the speech community of the targeted language or linguistic variety (cf. Biber, 1993, in Chapter 1).

Based on that, in Section 4.1, I will summarise the steps of this research, going through each Chapter and their respective Sections. Afterwards, Section 4.2 will account for the limitations of this study, what can be done in the future to improve or enhance its range, as well as which studies can be benefitted from it. It is worthy to have in mind not only the limitations, but the contributions.

4.1. Sum up

For the purpose of this research, this work was divided into four Chapters, as follows. Chapter 1 accounted for the theoretical framework. Two Sections, 1.1 and 1.2, introduced the general and the specific literature respectively. The former considered the interface between Corpus Linguistics and Historical Linguistics, since this study dealt with a past synchrony of English. This past synchrony is known nowadays as Middle English, comprehending the timespan between the 12th and 15th centuries. Two transitional periods are attested, the first one represented by the 12th century, or early Middle English after Old English. As for the

second, late Middle English was in process of giving place to Early Modern English, naturally the last century covered here, the 15th century. Any scholar of the history of English has notions on that, but for a lay audience it is more than necessary to mention these transitions in past synchronies of English without setting aside the abovementioned transitional periods.

Moreover, Corpus Linguistics had the most attention in a first moment in Section 1.1. Brilliant studies from old and modern scholars served as the anchor for this research, such as McEnery and Hardie (2011), stating the procedures followed by a Corpus linguist to analyse language or linguistic varieties. For the consolidation of this field of research in Linguistics, Kübler and Zinsmeister (2015) showed empirical methods for this type of analysis, more importantly the application of corpora "for testing their linguistic hypothesis" (p. 3). In addition, Weisser (2015) assumes that "a corpus is [...] based on a specific set of design criteria influenced by its purpose and scope" (p. 13). This author urges the linguist to be careful on what criteria should be more accurate to design a corpus and what for. The same author, as I will show further, has a website with a long list of historical corpora.

Another fact of importance is that, in present times, Corpus Linguistics demands machine-readable data. Being so, I cited two of the most relevant programming languages serving as sourcecodes for softwares, as known as R and Python. As for softwares used in linguistic analysis, *AntConc*, *TextSTAT*, *SketchEngine* and *UAM CorpusTool* are famous examples, among others.

Moving forward, still in Section 1.1, I presented the possibility for a not-yet-existing corpus to be of utmost interest for the linguist. In her study on parallel corpora, Lawson (2001) gave a brilliant solution for that in Ghadessy, Henry and Roseberry (2001) – "[i]f no suitable corpus is available, the alternative is to make one" (p. 293). This is a crucial notion

on the present study since no Middle English Corpus with free access was found to suit its needs, especially in the literary genre known as lyrics.

Accounting for core notions of Corpus Linguistics, size, representativeness and balance are regarded with a keener eye. Matching Lawson's assumptions on designing a corpus in the lack of an available one, the researcher's needs come in the first place. This is why Biber (1993) proved so important. In the case of this research, a small corpus is more than sufficient to analyse the occurrences of the Middle English verb *haven* ("to have") in a syntactic-semantic point of view. Following Ann Lawson, I compiled my own corpus, a convenience sample of philological editions of English lyrics in the centuries covered, named CMEL – *Corpus of Middle English Lyrics*. More details on that are described in Chapter 2.

Corpus Linguistics can serve as a "white flag" between fields, forming alliances with others. I cited some of these fields (e.g. Applied Linguistics, Pragmatics, etc), but the most important here is the one with Historical Linguistics. As one can see, although divergences between studies may arise, a machine-readable nature of a corpus seems to be a consensus. Corpora were given as examples of each of the abovementioned alliances, or here called interfaces. For Historical Linguistics, most of them accounting for past synchronies of English follow the skeletal structure of Penn-Treebank with its large number of parsed corpora (e.g. Nevalainen et al., 1998-2019; Zimmermann, 2014-2019 and others presented by Weisser (2009) in his website). Other languages such as both European and Brazilian Portuguese and Icelandic had their respective corpora as examples as well, most of them parsed.

Section 1.2 brought up the specific literature of this study concerning verb valency. As Linguistics longed for scientific status, a borrowing from a Chemistry terminology to Linguistics in the 20th century came well in hand. Then, a metaphor arises, comparing a verb to a molecule, and all the mandatory elements to reach stability in the molecule or the

sentence have to do with valency. In Linguistics, verb valency, considering the verb as the predicate. Important names such as Charles Peirce and Lucien Tesnière needed to be given attention. To ground this research, Perini (2015) was important too regarding verb valency. He presented a schema (p. 15) which I further adapted to the CMEL's sampled data of lexical occurrences of *haven*.

A question has arisen: *If changes in verb valency could be attested cross-linguistically, why not in the same language over centuries?* This question deserves some thinking and seems more than reasonable, given that my goal was to describe verb valency in each of the covered centuries and, in the end, compare what has changed or prevailed.

Chapter 2 concerned the methodological steps. Section 2.1 was dedicated to its description and compilation. The choice for timespan covered was justified. I brought forward difficulties faced in the compilation of this convenience sample of Middle English lyrics, and I will go back to them in Section 4.2. Also, the genre was introduced as well.

Following the theoretical framework presented in Chapter 1 and recapitulated in Section 4.1, the nature of data and decisions on text selection was essential. Then, I described how the data were processed in softwares after its careful selection from printed and online sources (cf. Wells, 1907; Hall, 1920; Brown, 1932, 1939; Mossé, 1968, and Duncan, 1995, all printed; and *TEAMS* project of the University of Rochester and the *Corpus of Middle English Prose and Verse* of the University of Michigan, both online). Softwares were handy in processing all data, some differences in the procedures for the two main retrieving sources. For instance, *ABBYY FineReader14* was necessary for optical recognition of image files in .png format resulting in a .txt file, for printed sources. Texts from online sources were copied and pasted in *Notepad++* and saved as .txt files. For all files in this format, although Corpus Linguistics defends nowadays the use machine-readable data, manual revisions were necessary. In the end, all the files were coded according to their centuries and text number in

order of compilation (see Table 1). The same was made for the metadata, but with the addition of "MD" (for "metadata") in the coded files.

Discussions about the manuscripts' philological editions were driven briefly. However, to avoid exhaustiveness through repetition, as the subcorpora for each century had different sizes, a sampling process through scripts in *R* software (R CORE, 2019) was of great need and help. This balancing process was detailed in Section 2.2.

Section 2.3 dealt with annotation. In other words, by means of *UAM CorpusTool*, version 3.3m, following Pustejovsky and Stubbs (2002). Serving as metadata about the texts, even before using this software, I gathered information about their manuscripts and editions as taking decisions on which ones would be preferable to be part of CMEL. Being so, knowledge of Middle English grammar was crucial, meaning fluency, taking into account the editors' notes as well as Mossé (1968). In the texts with any occurrence of *haven*, in the abovementioned software, besides carefully selecting the useful texts to fit this study's purposes, I translated the sentences where any form of *haven* took place. In *UAM CorpusTool*, which I classified through a schema all the occurrences having in mind that they are all context-based. This classification relied on their use as grammatical or lexical items, and, in the comments, my translation or few others (cf. Delamar, 2003; Thomas, 2008 and Treharne et al., 2010), changed when judged necessary. All this information is presented in Figure 6, Figure 7 and Figure 8.

The next step was creating spreadsheets – S01, S02 and S03. The first one accounted for the morphological occurrences of *haven* for each century, taking into account their glosses based on the Leipzig Glossing Rules (2015), before the classification. Appendices C1, C2, C3 and C4 illustrate these data. This, in turn, was done in S02, a new category inserted called "uncertain" (see Appendices D1, D2, D3 and D4). Their absolute and relative frequencies were shown in

Table **4** and Graph 1, for all centuries.

Section 2.4 was perhaps the most important one in Chapter 2. This is due to the sampling process, once again through *R* software (R CORE, 2019), to filter and sort lexical occurrences. These occurrences culminated in S03, the spreadsheet containing the syntactic-semantic analysis of each sorted occurrence in the samples for each century. An example is given in Table 5.

Then, Chapter 3 was subdivided into five Sections, the first four accounting for each century (3.1 for the 12th; 3.2 for the 13th; 3.3 for the 14th and 3.4 for the 15th). They follow the same structure, and over these I make brief comparisons an comments on morphological changes, influences from Old and Norman French (directly, due to the Norman Conquest in 1066 AD) and Classical Latin, the no-more spoken but still dominating language in all Europe culturally and religiously. The influence of these languages is remarkable for they play a very important role in linguistic change in English.

In the first four Sections of Chapter 3, Figures 9, 10, 11 and 12 showed the schema adapted from Perini (2015, see Figure 5 in Chapter 1) concerning verb valency. The most proper Middle English Dictionary judged here was the one developed by the University of Michigan. This choice has been justified over this thesis. Mainly, this source presents great information on morphology of *haven* and all the other words I had to look for in the translation process. In addition, it is corpus-based, so all the senses and subsenses of any word (let us give more attention to the verb *haven*, of course) are exemplified through editions of authentic texts. Its completeness influenced hugely on the choice of this dictionary over others. When mentioning *haven*'s etymology, the Online Etymology Dictionary (Harper, 2001-2019) proved its worth.

When talking about senses and subsenses, here they were called readings after Perini (2015). In a semantic perspective, this dictionary, here abbreviated as MED, was quite specific in showing a full range of meanings of *haven* in Middle English over its centuries. These senses and subsenses are coded, explained and exemplified. Accordingly, the readings received the same codes as well as their respective senses and subsenses. Appendices F1-A and F1-B (accounting for the 12th century), I1-A and I1-B (accounting for the 13th century), L1-A and L1-B (accounting for the 14th century) and O1-A and 01-B (accounting for the 15th century) showed the readings found for each century, their diatheses following Herbst et al.'s (2004) adapted model and then the syntactic orders for each. Collocations, figurative constructions, particular cases and relative clauses were marked in these Figures as well.

Frequency was relevant in this study. What was more frequent meant what was preferred by the Englishmen writing (or being the owner of the though behind the texts, in the case of apographs); what consisted of hapax legomena, meaning linguistic conservatism and/or some forms and meanings in the process of becoming old-fashioned. *What has changed?* Leading to *What was in the process of change over the centuries?* And also *What has prevailed?*. If a language only changes over time, there is no point on assuming that this is the same language or a mother-language that gave birth to daughter-languages, so, logically, something has to prevail – otherwise English would be nowadays called something else entirely. But a better question would be *Why?* In the end of all these. So I provided logical hypotheses to these questions and shed some light in the darkness of the reader's doubts.

Remarkably, in Section 3.2, the 13th century's sampled lexical occurrences showed a different behaviour compared to the other studied centuries. Its peculiarity lied on the nature of the corpus, namely a very big text representing around a half of the whole subcorpus. This caused a pattern shift in comparison to the others. The expected hypothesis for all centuries

was based on the process of grammaticalisation suffered by *haven* already taking place since Old English, when it began to be employed as an auxiliary verb in past tense formation and continued on until Present-Day English. Thus, only for this century's sample the relative frequency of lexical occurrences of this verb was higher than the one for grammatical occurrences, setting aside the "uncertain" occurrences. This affected all results for this century, giving rise to discussions. However, considering the four samples of CMEL diachronically, this peculiar behaviour did not impede a conclusion that the process of grammaticalisation was growing stronger.

Speaking of a given text influencing the whole subcorpus, the nature of the texts and thus the subcorpora must gain importance. In Section 3.5, I compared all the four centuries' results and was inclined to assume that the context is essential, as well as genre. Firstly, the 12th century had less variability of texts due to the difficulty of finding material of interest to be compiled. As for the 13th century, the same applies, but this time this subcorpus had a larger text which influenced in the whole analysis presented in Section 3.2, leading to a hypothesis that could not wholly influence the final results. In the case of the 14th century, more variability in texts' themes/subgenres was found, not forgetting the fact that some texts have known authorship, as well as the fact that no largest text was there to influence the analysis of this subcorpus. Finally, 15th century's subcorpus is a compilation of religious lyrics, so it was expected that some readings would prevail. In the end, all drives us to conclude a semantic analysis is context-based. Secondly, one must consider the literary genre here – lyrics. In spite of potential and linguistic freedom of poetic texts, it does not mean that the Englishmen organised their thoughts in spontaneous speech or when writing in other genres with such a syntactic flexibility. Probably one or two orders were preferred. As it has a musical nature, it is expected that metrics would be of great influence in the readings and in *haven*'s valency, and not be representative of spontaneous speech. As I have stated in this

study, an attempt to reconstruct historical spontaneous speech as well as its speech community is out of the scope here, and it would be illogical to get involved in such a quest using poetry as a base material. Once again, I call the reader's attention in what was said about decisions before designing a corpus and for what purpose.

As I mentioned in this sum up, Chapter 3 was subdivided into five Sections, but until now I have presented only the first four. A fifth, Section 3.5, served as an in-depth comparison between all studied centuries, focusing on the readings. The profile attested was similar for most centuries, setting aside the 13th, given its peculiar behaviour, but still with some similarities in the conclusions – such as the use of negative and irrealis (subjunctive) – but it is not the focus here. Keeping the focus on the readings, Figures 9, 10, 11 and were provided once more, and then I urged the reader to pay a very close attention to them. *Which readings are common to all centuries?*; *Which ones consist of innovations?*; *Which ones were peculiar to one specific century?*. These questions should be in his or her mind. Then, to answer, I organized these data in Appendix R, light purple and bold representing what remained common to all centuries, regardless of what could be more or less frequent. In other colours, I highlighted what was peculiar to one specific century (e.g. light green readings peculiar to the 12th century). Afterwards, I discussed these readings, focusing on the ones common to all centuries, resulting in Table 18 with their absolute and relative frequencies. Graph 10 followed to illustrate it. Comments were added regarding the senses and subsenses/readings (e.g. the incidence of group 4, about abstractness an ephemeral or intrinsic states or conditions).

Then, the valency received more attention, such as the common diatheses to all centuries and their readings. Appendices S1, S2, S3 and S4 show these data for each century, and finally

Table **19** shows the common argument structures for all centuries together. Four common diatheses, namely [N], [N PP], [N to-INF] and [N to-INF(haven)] were attested in all centuries. For [N], 'Arg1 V Arg2' prevails (three of seven orders) and has more variability of orders; [N PP] shown two prevailing orders in ten, as known as 'Arg1 V Arg2 Arg3' and 'Arg1 Arg3 V Arg2', the former with more occurrences. As opposed to this framework, the other two diatheses were less common and, naturally, with no common order for all centuries. Regarding Mossé (1968) in syntactic flexibility, he was not wrong. However, what I showed through

Table **19** was the *process* of syntactic fixity over the centuries covered, not the *result*; some orders, for the more common diatheses, were becoming regularised and others, the less frequent ones, were not. Naturally, Middle English, despite its morphological impoverishment in declensions and verbal inflections, was syntactically flexible and became fixed, but this *process* did not happen overnight, but over the centuries. Linguistic change takes time, and the most common phenomena are affected faster than the less common ones. This serves for the diatheses just presented.

After the sum up, I will help myself on presenting, in Section 4.2, the limitations of this study, how it can be improved and its contributions to further studies.

4.2. Further research

The way this research was conducted, every decision taken and method applied had a purpose: to investigate the syntactic-semantic behaviour of *haven*, by means of a corpus designed precisely for this study, namely CMEL – *Corpus of Middle English Lyrics*. More specifically, the interpretation of the written production on the chosen genre was greatly important to infer its senses as subsenses, as the Middle English Dictionary developed by the

University of Michigan presented. Drawing on Perini (2015), these senses and subsenses were called readings.

This research followed a robust theoretical framework, cited in the previous Section, followed procedural steps and reached remarkable results. Section 4.1 served to summarise all what has been done so far, and yet this study may present limitations and difficulties. This is not bad. Indeed, I see it positively since perfection is merely a myth, something that nothing and no one can achieve, yet can give us good lessons.

One of these lessons is how to improve this study and overcome the difficulties and limitations. I recognise that this is a Master's thesis, with much yet to be learned. More accurate knowledge on Middle English is one of the best examples alongside mastery on an analytical process of syntactic-semantic behaviour of any word, here the verb *haven*. For instance, my translations can be further edited and improved. In addition, programming skills to optimise the task could have been better used, and also the softwares mentioned over this thesis, especially automatising the classification of the forms of *haven* in *UAM CorpusTool*, version 3.3m, instead of setting it to manual annotation. These are my personal limitations.

Not only personal, but theoretical and methodological limitations exist as well. For instance, I cannot ignore the fact that I am part of the Brazilian academic community, which has little or no interest in researching the history of a Germanic language – in this case, English. Unfortunately, material to compile a corpus of any kind regarding Middle English are scarce in Brazil and even online for free or with a cheap price. This is the justification on why I designed a convenience sample, having on hand what was available and employing such a great effort to find not only texts, but also theoretical works, with the help of students overseas. With a greater availability of material to compile a corpus and to anchor such a study, it could be indeed enhanced, more robust. The issue on the 13th century, for instance, could be avoided with this solution.

Accounting for the nature of all subcorpora, a more precise process of balancing/sampling can be conducted in a more varied, stratified corpus, not only of lyrics, but also contemplating another written genre(s) with a broader range of uses. In other words, CMEL someday can become more judicious and varied. In this research, I focused on syntactic-semantic behaviour of *haven*, more specifically verb valency when employed as a lexical item, but further researches with an enhanced corpus can undoubtedly be handy to study other verbs, or even other word classes and sentences as a whole. Besides, other domains, such as involving extralinguistic factors (Corpus Linguistics in an interface with Historical Sociolinguistics) should be explored as well, as, in fact, it already is (cf. Nevalainen and Raumolin-Brunberg, 2003, for Early Modern English, among others).

Despite of all limitations faced, based on my corpus and my work as a whole, other studies can be benefitted. I call the attention to future researchers in the Brazilian academic community to become curious about the history of English and have some insight to keep the flow and contribute to the academic production on the study of past synchrony of Germanic languages. My corpus is free of charge, and I would be more than satisfied if someone used it accurately to conduct further studies on Corpus Linguistics in interface with a variety of fields. I will give examples through questions: a) Applied Linguistics: *how could knowing a past synchrony of English help English learners?* b) Translation Studies: *how could a Middle English corpus and a thesis based on it contribute to translation methods in poetry?*; c) Lexicography and spelling variation: *why do some registered word forms for the same gloss prevail as opposed to others? And how about linguistic policy and dictionarised forms?*; d) Music and Literature Studies: *how can the metrics and melody of a poem to be sung on be reconstructed or adapted to give us a glance of how medieval English music used to sound?*, among others.

I gave some ideas. Now, it is time for future researchers to develop them or think about their own.

REFERENCES

ABBYY (2013). ABBYY FineReader (Version 14).) [Computer Software]. Retrieved from https://www.abbyy.com

Anthony, L. (2019). AntConc (Version 3.5.8) [Computer Software]. Tokyo, Japan: Waseda University. Retrieved from https://www.laurenceanthony.net/software

Biber, D. E. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, 8(4), 243-257. https://doi.org/10.1093/llc/8.4.243

Brooklyn-Geneva-Amsterdam-Helsinki Parsed Corpus of Old English. (2003). Retrieved December 07, 2019, from http://www-users.york.ac.uk/~sp20/corpus.html

Brown, C. (ed.) (1932*). English Lyrics of the XIIIth century*. Oxford, Oxfordshire: Oxford University Press.

Brown, C. (1939). *Religious lyrics of the XVth century*. Oxford, Oxfordshire: Clarendon.

Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The evolution of grammar: tense, aspect and modality in the languages of the world*. Chicago; London: The University of Chicago Press.

Bynon, T. (1977/2012). *Historical Linguistics*. Cambridge, Cambridgeshire: Cambridge University Press. https://doi.org/10.1017/CBO9781139165709

Camargo, D. C. D. (2008, July). *Pesquisas em Tradução e Lingüística de Corpus.* Paper presented on the meeting of ENANPOLL – XXIII Encontro Nacional da ANPOLL, São José do Rio Preto, São Paulo, Brazil. Retrieved from <http://letra.letras.ufmg.br/gttrad/enanpoll/2008/9%20-%20Pesquisas%20em%20tradu%C3%A7%C3%A3o%20e%20ling%C3%BC%C3%ADstica%20de%20corpus.pdf>.

Castro, R. A. (2016*). Desenvolvimento, Implementação e Teste de Ferramentas Integradas para Análise Textual e Tratamento Estatístico de Dados em Pesquisas Linguísticas*. (Master's thesis, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil). Retrieved from http://www.poslin.letras.ufmg.br/defesas/1741M.pdf

Cavalcante, F. A. (2015). *The topic unit in spontaneous American English: a corpus-based study*. (Master's thesis, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil). Retrieved from http://hdl.handle.net/1843/MGSS-A7GQ48 Accessed on 2019-12-07.

CLP – Corpus Lexicográfico do Português. (2017). Retrieved December 07, 2019, from http://clp.dlc.ua.pt/Inicio.aspx

Comrie, B. (2006). Transitivity pairs, markedness, and diachronic stability. *Linguistics*, 44(2). 303-318. https://doi.org/10.1515/LING.2006.011

C-ORAL Brasil Project – Corpus de referência do português brasileiro falado informal. (2012). Retrieved December 07, 2019, from http://www.c-oral-brasil.org/

Corpus Compartilhado Diacrônico: Cartas pessoais brasileiras. (n.d.). Retrieved December 07, 2019, from http://www.tycho.iel.unicamp.br/laborhistorico/corpora.html

Cresti, E. *Corpus di Italiano Parlato*. (2000). Firenze, Italy: Accademia della Crusca. v. I-II [CD-ROM]

Delamar, M. H. *"Truste e raþir to lettirs written in þ'is": A Selection of Middle English Death Lyrics in Translation*. (Master's thesis, North Carolina State University, Raleigh, North Carolina, United States). Retrieved from https://repository.lib.ncsu.edu/bitstream/handle/1840.16/1689/etd.pdf?sequence=1&isAllowed=y

Dryer, M. S. & Haspelmath, M. (eds.) 2013. The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved December07, 2019, from http://wals.info

Duncan, T. G. (1995). *Medieval English Lyrics: 1200-1400*. London; New York: Penguin Books.

Eckardt, R. (2006). *Meaning change in grammaticalization: An enquiry into semantic reanalysis*. New York, United States: Oxford University Press.

Galves, C., Andrade, A. L. D., & Faria, P. (2017, December). *Tycho Brahe Parsed Corpus of Historical Portuguese*. Retrieved December 07, 2019, from http://www.tycho.iel.unicamp.br/~tycho/corpus/texts/psd.zip

Ghadessy, M., Henry, A., & Roseberry, R. L. (2002). *Small Corpus Studies and ELT: Theory and practice*. Amsterdam; Philadelphia: John Benjamins Publishing Company.

Giancarlo, M. (2001). The Rise and Fall of the Great Vowel Shift? The Changing Ideological Intersections of Philology, Historical Linguistics, and Literary History. *Representations, 76,* 27-60. Retrieved from https://www.academia.edu/12530905/The_Rise_and_the_Fall_of_the_Great_Vowel_Shift_The_Changing_Ideological_Intersections_of_Philology_Historical_Linguistics_and_Literary_ History

Gries, S. Th. (2009). *Quantitative Corpus Linguistics with R: A Practical Introduction*. Abingdon, Oxfordshire; New York, New York: Taylor & Francis Group.

Hall, J. (1920). *Selections from the Early Middle English. 1130-1250*. 2 parts. Oxford, Oxfordshire: Clarendon.

Harper, D. *Online Etymology Dictionary*. Retrieved December 07, 2019, from https://www.etymonline.com/

Henry, A., & Roseberry, R. L. (2001). Using a small corpus to obtain data for teaching a genre. In: Ghadessy, M., Henry, A., & Roseberry, R. L. (2002). *Small Corpus Studies and ELT: Theory and practice*. Amsterdam; Philadelphia: John Benjamins Publishing Company. 93-133.

Herbst, T., Heath, D., Roe, I. F., & Götz, D. (2004). *A valency dictionary of English: a corpus-based analysis of the complementation patterns of English verbs, nouns and adjectives*. Berlin; New York: Mouton de Gruyter.

Ho, D. (2003-ongoing). Notepad++ [Source code editor]. Retrieved from https://notepad- plus-plus.org/

Hogg, R. (2002). *An Introduction to Old English*. Edinburgh, Scotland: Edinburgh University Press Ltd.

Jargas, A. M. (2012). *Expressões regulares: uma abordagem divertida*. (4th ed.). São Paulo, Brazil: Novatec Editora Ltda.

Kossel, W. Über Molekülbindung als Frage des Atombaus. *Annalen der Physik*, 4(49). Retrieved December 07, 2019, from https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19163540302

Kübler, S., & Zinsmeister, H. (2015). *Corpus Linguistics and Linguistically Annotated Corpora*. London, England: Bloomsbury.

Kurylowicz, J. (1965/1975). The Evolution of grammatical theories. *Diogenes*, 51. 55-71.

Kytö, M. (2011). Corpora and Historical Linguistics. *Revista Brasileira de Linguística Aplicada*, 11(2). http://dx.doi.org/10.1590/S1984-63982011000200007

Lamont, G. (2005). *The Progress of English Verb Tenses and the English Progressive*. Retrieved on December 07, 2012, from http://homes.chass.utoronto.ca/~cpercy/courses/6362- lamont.htm

Lawson, A. (2002). Collecting, aligning and analysing parallel corpora. In: Ghadessy, M.; Henry, A.; Roseberry, R. L. (2002). *Small Corpus Studies and ELT: Theory and practice*. Amsterdam; Philadelphia: John Benjamins Publishing Company. 279-309.

The Leipzig Glossing Rules: Conventions for interlinear morpheme-by-morpheme glosses. Department of Linguistics: Max Plank Institute. Retrieved December 7, 2019, from https://www.eva.mpg.de/lingua/resources/glossing-rules.php

Lexical Computing CZ s.r.o. (n.d.). *SketchEngine*. Retrieved December 07, 2019, from https://www.sketchengine.eu/

Lewis, G. N. (1916). The Atom and the Molecule. *Journal of the American Chemical Society*, 38. 762-785. Retrieved from https://pubs.acs.org/doi/10.1021/ja02261a002. https://doi.org/10.1021/ja02261a002

Marcus, M. P., Santorini, B., & Marcinkiewicz, M. A. (1993). *Building a Large Annotated Corpus of English: the Penn Treebank*. Pennsylvania, United States: University of Pennsylvania.

McEnery, T., & Hardie, A. (2011). *Corpus Linguistics: Method, Theory and Practice*. Cambridge, Cambridgeshire: Cambridge University Press.

Middle English Compendium (2018). *Middle English Dictionary*. Retrieved December 07, 2019, from https://quod.lib.umich.edu/m/middle-english-dictionary/dictionary

Mossé, F. (1968). *Handbook of Middle English*. Baltimore, Maryland: The Johns Hopkins Press.

Nevalainen, T., & Raumolin-Brunberg, H. (1993-ongoing) Corpus of Early English Correspondence. Retrieved December 07, 2019, from http://www.helsinki.fi/varieng/CoRD/corpora/CEEC/index.html

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

Nichols, J., Peterson, D., & Barnes, J. (2004). Transitivizing and detransitivizing languages. *Linguistic Typology*, 8(2). 149-211.

O'Donnel, M. (2017). UAM CorpusTool (Version 3.3m) [Computer Software]. Retrieved from http://www.corpustool.com/

Pagliuca, W. (1994). *Perspectives on Grammaticalization*. Amsterdam, Netherlands: John Benjamins.

Partington, J. R. (1921). *A Text-book of Inorganic Chemistry for University Students*. London, England: Macmillan and Co., Limited.

Penn-Helsinki Parsed Corpus of Early Modern English. (n.d.) Retrieved December 07, 2019, from https://www.ling.upenn.edu/hist-corpora/PPCEME-RELEASE-3/index.html

Penn-Helsinki Parsed Corpus of Historical English. (n.d.). Retrieved December 07, 2019, from https://www.ling.upenn.edu/hist-corpora/

Penn-Helsinki Parsed Corpus of Middle English, 2nd edition, release 4. (n.d.). Retrieved December 07, 2019, from https://www.ling.upenn.edu/hist-corpora/PPCME2-RELEASE-4/index.html

Penn Parsed Corpus of Modern British English, 2nd edition, release 1. (n.d.). Retrieved December 07, 2019, from https://www.ling.upenn.edu/hist-corpora/PPCMBE2-RELEASE-1/index.html

Perini, M. A. (2015). *Describing verb valency: Practical and theoretical issues*. London, England: Springer.

Pustejovsky, J., & Stubbs, A. (2002). *Natural Language Annotation for Machine Learning: A Guide for Corpus-building Applications*. Cambridge, Cambridgeshire: O'Reilly Media.

Python Core Team (2015). Python: A dynamic, open source programming language. Python Software Foundation. Retrieved from https://www.python.org/.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/.

Ramat, P. (2012). The (Early) History of Linguistic Typology. *The Oxford Handbook of Linguistic Typology*. 1-10. 10.1093/oxfordhb/9780199281251.013.0002

Raso, T., & Mello, H. (2009). Parâmetros de compilação de um corpus oral: o caso do C- ORAL Brasil. *Veredas On-line*, 13(2), 20-35. Retrieved December 07, 2019, from http://www.ufjf.br/revistaveredas/files/2009/11/ARTIGO-Tommaso-Raso-e-Heliana- Mello.pdf

Rocha, C. F. (2014). A coleta de corpus de aprendizes questões qualitativas em uma pesquisa sobre a escrita de aprendizes de língua espanhola. *Estudos Linguísticos*, 43(1), 286-297. Retrieved December 07, 2019, from https://revistas.gel.org.br/estudos-linguisticos/article/view/438

Shields, K. (2012). Linguistic Typology and Historical Linguistics. *The Oxford Handbook of Linguistic Typology*. 1-12. 10.1093/oxfordhb/9780199281251.013.0026

TEAMS - Middle English Texts Series Online. (s.d.). Robins Library Digital Projects. Retrieved from https://d.lib.rochester.edu/teams/text-online Accessed on 2019-12-07.

Tesnière, L. (2015). Elements of Structural Syntax. In Osborne, T., Kahane, S. (Trans.), *Elements of Structural Syntax*. (Original work published 1959).

TextSTAT (2015) [Computer software]. Retrieved from http://neon.niederlandistik.fu- berlin.de/en/textstat/

Thomas, C. M. (2008). *"Poema Morale": An Edition from Cambridge, Trinity College B. 14.52.* Retrieved December 07, 2019, from http://purl.flvc.org/fsu/fd/FSU_migr_etd-1573

Toyota, J. (2002). Lexical analysis of Middle English passive constructions. In Díaz-Vera, J. E. (ed.) (2002). *A Changing World of Words: Studies in English Historical Lexicography, Lexicology and Semantics*. Amsterdam; New York: Rodopi B. V. 572-610.

Treharne, E., Walker, G., & Green, W. (eds.) (2010*). The Oxford Handbook of Medieval Literature of English*. Oxford, Oxfordshire: Oxford University Press.

van Gelderen, E. (2005). *Structures behind grammaticalization*. Oslo; Arizona: Arizona State University.

van Gelderen, E. (2011). Valency changes in the history of English. *Journal of Historical Linguistics*, 1(1). 106-143.

Visser, F. (1970). *An Historical Syntax of the English language*. Leiden, Netherlands: Brill.

Wallenberg, J. C., Ingason, A. K., Sigurðsson E. F., & Rögnvaldsson, E. (2011). *Icelandic Parsed Historical Corpus (IcePaHC)*. Version 0.9. Retrieved December 07, 2019, from http://www.linguist.is/icelandic_treebank

Weisser, M. (2016-2019). *Historical & Diachronic Corpora*. [HTML webpage] Retrieved December 07, 2019, from http://martinweisser.org/corpora_site/historical_corpora.html

Wells, J. E. (ed.) (1907). *The Owl and the Nightingale*. Boston, Massachusetts: D. C. HEATH AND CO.

Zimmermann, R. (2014-2019). *Parsed Corpus of Middle English Poetry*. Retrieved December 07, 2019, from https://pcmep.net/

APPENDICES

A1 – Script for sampling CMEL's 12th century texts.

```
# Sorting CMEL's texts - 12th century

############# Preliminary steps
# show objects in the memory
ls()
# clean memory
rm (list=ls())
# check current directory
getwd()
# save directory in an object
maindir = getwd()
# save directory in a secondary object
secDir = setwd(choose.dir())
# show directory files
#dir()

# sampling
# total number of texts
#n = length(dir())
n = 32

# sample size
nsample = 4

# sorting
set.seed(455)
sample1 = sample(n,nsample)
sample1

# sampled texts
#dir()[sample1]

file1 = read.csv2(choose.files())
file1
str(file1)
file1[sample1,]
new_file = file1[sample1,]
new_file
sum(file1[sample1,2])
write.csv2(new_file,"sample-sec12.csv")

# needed files
new_file2 = file1[-sample1,]
new_file2

write.csv2(new_file2,"sample-neededfiles12.csv")
```

A2 - Script for sampling CMEL's 13th century texts.

```
# Sorting CMEL's texts - 13th century

############# Preliminary steps
# show objects in the memory
ls()
# clean memory
rm (list=ls())
# check current directory
getwd()
# save directory in an object
maindir = getwd()
# save directory in a secondary object
secDir = setwd(choose.dir())
# show directory files
#dir()

# sampling
# total number of texts
#n = length(dir())
n = 71

# sample size
nsample = 46

# sorting
set.seed(455)
sample1 = sample(n,nsample)
sample1

# sampled texts
#dir()[sample1]
file1 = read.csv2(choose.files())
file1
str(file1)
file1[sample1,]
new_file = file1[sample1,]
new_file
sum(file1[sample1,2])
write.csv2(new_file,"sample-sec13.csv")

# needed files
new_file2 = file1[-sample1,]
new_file2

write.csv2(new_file2,"sample-neededfiles13.csv")
```

A3 - Script for sampling CMEL's 13th century texts.

```
# Sorting CMEL's texts - 14th century

############# Preliminary steps
# show objects in the memory
ls()
# clean memory
rm (list=ls())
# check current directory
getwd()
# save directory in an object
maindir = getwd()
# save directory in a secondary object
secDir = setwd(choose.dir())
# show directory files
#dir()

# sampling
# total number of texts
#n = length(dir())
n = 54

# sample size
nsample = 3


# sorting
set.seed(455)
sample1 = sample(n,nsample)
sample1

# sampled texts
#dir()[sample1]
file1 = read.csv2(choose.files())
file1
str(file1)
file1[sample1,]
new_file = file1[sample1,]
new_file
sum(file1[sample1,2])
write.csv2(new_file,"sample-sec14.csv")

# needed files
new_file2 = file1[-sample1,]
new_file2

write.csv2(new_file2,"sample-neededfiles14.csv")
```

B – MED's enlisted uses and meanings for the entry 'haven v.'.

hăven v.

Entry Info

Forms hăven v. Also *have, ave(n, haf(e(n, haffe, haif, haw(e(n, (early) hafven, hafa*, (late) *ihave,* (error) *hove & habbe(n, abbe(n, habe(n, h)ab,* (early) *habban, habbæn, habbeon, ȝehabban, hebbe, hæbbe(n* & *han, an, hanne, ha, a, hai* & *halven, half(e;* neg. *nave, nabbe(n, næbbe.* Forms: sg. 1 *hăve,* etc. & *hăs;* neg. *năve, năf(e, nabbe;* sg. 2 *h)ăvest, h)ăfest, hăves, hăvez, hafs,* (early) *hafast, hafust, hævest, hafvest, hæfvest, hafst* & *habbes, habbez* & *h)ăst, haist, hĕst, hăs(e, ăs, hasse, hătz* & (early) *hæhvest*; neg. *năvest, nafst, nefst, năst;* sg. 3 *h)ăveth, hăved, h)ăvet, h)ăfeth, hăveht, h)ăves, hăvez, hafs,* (early) *hăfað, hafæð, hæveð, hæfeð, hæfed, hæfet, hæfveð, hæfæð, hafð, hafd, hefð, hæfð, hæfd, hafh* & *habbes, habbez* & *h)ăth, hăd, haith, haeth, hăt, hĕth, hĕt, haght, haȝt, haȝth, hăs(e, ăs, hasse, haes, hătz* & (?errors) *hatȝ, haȝ, heȝ;* neg. *năveth, năfeth, năved, năvet* & (early) *nafæð, nafð, næfð, năbit;* pl. (including impv.) *hăv(e(n, hăf(e(n, haffe, hăveth, hăves, hăfez* & *habben, haben, h)abbeth, habeth, h)abbet, habbez,* (early) *habbaþ, habbæð, habbeoð* & *h)ăn, hă, a, hăth, haght, haht, hăs(e, ăs, hais, hătz,* (errors) *hathes, hases* & *half(e, helveþ; neg. năve, nabbe, nabbeth, nabbed, nabbet,* (early) *nabbæð, nabbeoð, næbbæð;* ppl. *hăving(e,* etc. & (pl.) *awends;* p. *h)ăved(e, hevede, (early) hafede, hafvede, hefede, hefvede, hævede, hæfede, hæfvede, heovede & havde, hafd(e, afde,* (early) *h)efde, hæfde, h)eafde, heofde, hafte, hefte* & *h)adde, hădd(e, ad, haid(e, hedde, hĕd(e,* (early) *hædde* & (early) *hehde, hæhde, hæhvede;* neg. *nadde, năd(e, nhadde, nedde,* (early) *nevede, nefede, nævede, nafde, nefde, næfde, neafde;* sg. 2 *haddest,* etc. & *hădes, hădez,* (early) *hafdis;* neg. *naddest,* etc.; ppl. *i)hăved, ăved, ihăvet, h)ihafd, ihĕved* & *i)hădd, i)hadde, jhădd, hădde, i)hĕd, ihĕt,* (early) *ihæd.* Contractions: inf. *tave* (**= to have**), *tan, ta, taben* (= to have ben), *taclipsed, tadaued, tadwelled, tafalle, talived, hant* (**= han it**); sg. 1 *ichave, ichabbe, ic(h)chabbe, havi, habbich;* neg. *inabbe, navi, nabbi, nab(b)ich;* sg. 2 *havestu, hastou;* neg. *navestu, nevestu, nastou;* pl. *yave* (**= ye have**), *houwe* **(?= have we**), *havt* **(= have it**); p. *hefdich, haddich, had(d)estou, haddestu;* neg. *neddi, nefdich, nadestou, nede* (**= ne hadde he**), *nadda* (**= ne hadde heo**).

Etymology OE **ge)habban**; sg. 2 **hafast** (Nhb. **hæfes**, WS **hæfst**); sg. 3 **hafaþ** (Nhb. **hæfes**, WS **hæfþ**); pl. **habbaþ** (Nhb. **habbas**); p. **hæfde** (Merc. **hefde**); ppl. **hæfd** (Merc. **hefd**); neg. **næfst, næfeþ, nabbaþ, næfde**, etc. Forms with **-lf-**, **-lv**-are reverse spellings which reflect the weakened pronunciation of l before the labiodental; cp. **haf**, var. of **half**, etc. Forms like **hehde, hæhvede** in Layamon show an orthographic substitution of **h** for **f**; later forms **haght, haht, haȝt**, etc., are also orthographic. Past forms in present conditional or subjunctive senses appear in the 12th century, become frequent in the 14th.

Definitions (Senses and Subsenses)

1a. (a) To possess (sth.), own; **nadda non**, she had none; **we havt**, we have it; ppl. **havand**, wealthy; (b) to possess (land, an estate or inheritance, a city, town, manor, etc.); -- also used *fig.*; (c) to rule (a kingdom or part of a kingdom); (d) to inhabit (a country, house, temple, etc.); (e) **~ in (on, an) hond**, **~ in governaunce (governinge, possessioun, weldinge)**, **~ on anwald (anwealdnesse, weld, aughte, baillie)**, to have (sth.) in (one's) possession or under (one's) control; -- with or without gen. denoting possessor; **~ in heritage,** possess (land) by hereditary right; **~ in havendes (awends)**, have (sth.) in (one's) possession or keeping; (f) in legal phrases: **to ~ and to holden**, to possess and retain possession; **to ~ and to holden to,** to be had and retained by (sb.); for to ~ to, to be possessed by (sb.); **to ~ and enjoien (rejoicen)**, **to ~ usen and holden**, etc.; (g) in proverbs: ppl. **havande**, wealth.

1b. (a) To have (sb. or sth. somewhere); have (sth. in a position or direction); **~ mid**, have (sb.) with (one); **~ to (til) hond**, have (sth.) at hand; (b) to have (sth. in a book or document).

1c. With object and infinitive: to possess (sth. together with the capacity to do sth. with it); **~ to yeven**, have (sth.) to give, etc.; -- also with obj. unexpressed.

1d. (a) To have (sb. or sth. in a state or condition); ~ **in blis**, have (sb.) in bliss; ~ **in custume (usage, wone)**, be in the habit of (doing sth.), be accustomed (to do sth.), do (sth.) habitually; ~ **in doute**, fear or dread (sth.); ~ **in meditacioun**, meditate upon (sth.); ~ **in mouth**, have (sth.) in one's speech, talk about (sb. or sth.); ~ **in propre**, have personal possessions; ~ **in worshipe**, worship (God); ~ **on use**, be accustomed to say (a word); (b) ~ **in (on) hond**, ~ **in ward,** ~ **under cure**, to have charge of (sb. or sth.); (c) ~ **in (on) hond**, to be concerned with (sb. or sth.), be busy about, pay attention to.

2. (a) Of living things: to be provided with (a part, an organ); have (a part or an organ of a certain kind); have (a soul, body); have (a growth on some part of the body); (b) in *fig.* phrases: ~ **herte**, to have the heart (to do sth.), bear (to do sth.); have a desire (to do sth.); ~ **on herte**, be in agreement, be of one mind; ~ **gret (hol, mannes, reuful) herte**; ~ **tonge**, be able to speak; have a tongue (able to say sth.); ~ **eie,** ~ **fot,** ~ **hond**; -- also in prov.; (c) to wear (clothing, armor); ~ **on (upon)**, have on (a garment, etc.); ~ **on hond**, wear (a ring) on one's hand; (d) of a house, wagon, city, coin, etc: to be furnished with (sth.); of a country: have (rivers); of a pot: contain (sth.); of a roast goose: be accompanied by (wine); of a Gospel passage: be provided with (an exposition).

3. (a) To have (sb.) under one; have (a servant, slave, attendant, soldier, subordinate); command (an army or part of an army); ~ **and holden**, ~ **under**; (b) to have (sb.) in custody; hold (a prisoner); of devils: have power over (sb.); of disease: grip (sb.); ~ **on hond**; (c) to have (sb. in a certain relationship); have (a debtor, friend, physician, king, comrade, teacher, an enemy, etc.); have (God, the Holy Ghost, a god); ~ **to ifere**, have (sb.) as a companion; ~ **to maister**, have (sb.) as a teacher; (d) to have (a father, mother, son, daughter, husband, wife, etc.); ~ **to doughter**, have (sb.) for a daughter; ~ **to quene**, ~ **to sone;** ~ **to (unto) wif**, be married to (a woman); (e) ~ **evening (felaue, per)**, to have an equal; (f) ~ **on fore-ward**, *refl.* to be bound by an agreement.

4a. To have (an abstraction): (a) to enjoy (a right or privilege), have (the right to possess sth.); have (power, jurisdiction, authority); hold (an office); have (leave or license to do sth.); ~ **titel**, have a just claim to something; (b) ~ **name (note)**, to have a name; also, have a (certain kind of) reputation; (c) ~ **charge** (cure, hed, ward), to have charge (of sth.); ~ **correccioun**, have power to punish or discipline; (d) to have (witness, testimony, evidence, proof, an example); (e) to have (a law, commandment); be bound by (an oath); have (an agreement with sb.); **havinge a vou on hem**, being under a vow; (f) to have (time, a certain amount of time, enough time, a short time); ~ **time,** ~ **space,** ~ **while;** ~ **tom (leiser)**, have leisure; ~ **houre**, have a fixed time; ~ **time and stede**, have an appropriate time and place; (g) ~ **cure**, of a disease: to have a cure or treatment; ~ **hevi cariage**, of a cart: be hard to pull; ~ **place**, of a medicine: have a proper time to be used; ~ **times**, of a disease: have stages; (h) ~ **part in (on),** ~ **ani thing in**, to have power over (sb. or sth.); ~ **part (parti) of**, have anything to do with (sb. or sth.).

4b. To be in a state or condition: (a) to have (happiness, freedom, honor, salvation, grief, hardship, need, strife, etc.); ~ **scarsete**, be scarce; ~ **respect to**, be connected with, have a relationship to; (b) to have (a disease, vermin), suffer (pain, harm, a blow, a wound); be possessed by (devils); of a gem: have (a blemish); ~ **cold**, be cold; (c) to have (cause or reason), have (reason to do sth.); have (a complaint or grievance against sb.); (d) ~ **nede of (to)**, to need (sth.); ~ **nede that**, need or require that (sth. be or happen); (e) ~ **nede**, to be needy, be in want; (f) ~ **nede**, to need (to do sth.); be under compulsion, be in trouble, need help; (g) ~ **inough**, to have enough, call it quits; with inf.: be barely able (to do sth.); ~ **inough to don (beren, turnen)**; (h) **what hast thou**, what is the matter with you?

4c. To possess a quality or attribute: (a) to have (a nature, quality, characteristic, etc.); ~ **kinde**, have the nature (of sth.); ~ **token**, have a characteristic; ~ **tokeninge (bitokninge)**, have a significance; ~ **wit**, have intelligence; ~ **minde,** ~ **wille**, etc.; (b) to be endowed with (a physical quality); have (beauty, size, brightness, heat, a color, taste, voice, etc.); (c) to have (a function, power, etc.); ~ **lif**, have life, be alive; ~ **sight**, be able to see; ~ **strengthe**, have strength; also, have a medicinal property; (d) to have (a moral or spiritual quality); have (a virtue, vice); have (a certain kind of manners or bearing); ~ **sinne (gilt)**, have sin (guilt), be sinful (guilty); (e) to show (a quality); ~ **merci of (on, upon, to),** ~ **milce of,** ~ **ore of,** ~ **pite of (on, upon),** ~ **reuth of (on)**, show mercy to (sb. or sth.), take pity on; ta merci, to have mercy; (f) to be of (a certain age), be (a certain number of years old); ~ **fifti yer**, be fifty years old; ~ **age**, be of age; ~ **non age**, be too young; (g) ~ **right**, to be right, be in the right; ~ **wrong (unright, wough)**, be wrong, be in the wrong, be mistaken; (h) ~ **English**, to have the ability to express (sth.) in English.

4d. To have (an emotion, idea, etc.), have (sth.) in the mind or heart: (a) to experience (a feeling, fear, pleasure, wonder, joy, etc.); have (will, desire, envy, etc.); ~ **pite**, feel pity; (b) to have (an idea, a thought), have (knowledge, etc.); know (a name, terms, that sth. will happen); **nede (=ne hadde he)**, he had nothing, he knew of nothing [quot.: a1325]; (c) ~ **minde**, to remember (sth.), consider; ~ **remembraunce**, remember (to do sth.); ~ **iminde of**, have interest in (sth.); ~ **minde of (on, upon),** ~ **minning o**, remember (sb. or sth.); be mindful of, consider; ~ **remembraunce of**, remember (sth.); (d) ~ **reward to (unto, at),** ~ **biholding to,** ~ **resport to,** ~ **ward o**, to take thought of (sth.), have regard for; (e) ~ **in herte (wit)**, to feel (sth.) in the heart (mind); ~ **in minde,** ~ **on herte**, intend or desire (sth.) in the mind (heart); ~ **in knouleche**, know (sth.); ~ **in wille**, intend (to do sth.).

5a. (a) To hold (sth.), keep, grasp; retain (one's food); ~ **in (on)**, hold (sth.) in (one's mouth, one's hand or hands); ~ **on**, keep (sth.) on (one's person); ~ **abouten,** ~ **bi the throte,** ~ **mid**; (b) to keep (an idea, emotion, a memory) in (one's mind, heart), have (sb.) in (one's mind or heart), remember (sb. or sth.); ~ **in minde (herte, thank, thought)**; ~ **in memorie**, remember (sb. or sth.); (c) to have (a habit or custom), follow (a practice); (d) to preserve (sb. or sth.), save; keep (life, limbs); protect (sb. from sth.); God have you, God keep you, etc.; ~ **and holden**; (e) *refl.* to maintain oneself, conduct oneself, behave; of the stomach: behave, react; (f) ~ **pes**, to hold (one's) peace, be silent.

5b. (a) To keep (sb. or sth. in a place); fig. ~ **with him**, have (sth.) in his favor; (b) to keep (sb. or sth. in a state or condition); ~ **redi (yare)**, keep (sb. or sth.) in readiness; ~ **clene**, keep (one's body) clean; (c) to direct (sth.), incline; ~ **on**, keep (one's thoughts) on (sb.), have (one's heart) in (an activity); ~ **to (til)**, keep (one's eye) on (sth.), incline (one's heart) to (sb.), fix (one's hope) on (sth.); ~ **toward**, direct (the eye) toward (sth.); ~ **upwardes**, direct (one's eyes) upwards.

6a. (a) To regard (sb. or sth. in a certain way), consider; ~ **dere**, hold (sb. or sth.) dear, regard highly; ~ **suspect**, suspect (sb.); (b) ~ **excused**, to excuse (sb.), forgive, pardon; (c) ~ **in**, to hold (sb. or sth.) in (a certain estimation); ~ **in hate**, hate (sb.); ~ **in suspect**, be suspicious of (sth.); -- with noun; ~ **as for nought**, regard (sb.) as nothing, hold in contempt; ~ **on imis**, take (sth.) amiss, misunderstand; ~ **to hate**, hate (sb.); ~ **to hosp**, blaspheme (the Holy Ghost); ~ **as**, regard (sb. or sth.) as; -- with noun or adj.; ~ **for**, regard (sb.) as; -- with adj.

6b. In expressions of comparison or preference: (a) ~ **as lef..as**, would as willingly have (sth.) as (sth. else); would as soon (be or suffer sth.) as (do, be, or suffer sth. else); ~ **as lef**, would be as well pleased (if sb. did sth.); -- with noun, inf., or clause; (b) ~ **lever..than**, would rather have (sb. or sth.) than (sth.else); would rather have (sth.) than (be, have, or suffer sth.); -- also refl.; (c) ~ **lever..than**, would rather (do or suffer sth.) than (sth. else); ~ **lever**, would rather (do, have, or suffer sth.), would prefer; -- also refl.; (d) ~ **lever..than**, would rather (that sth. happened or were true) than (sth.); (e) ~ **levest**, would most prefer; -- with inf.; (f) ~ **rather**, would prefer (to do or have sth., that sth. should happen), would rather have (done sth.); -- with inf., ppl., or clause.

7a. (a) To obtain (sth.), take, get, receive; ~ **up**, have all of (sth.); **hant**, have it; [see also **onyen** ~]; (b) to capture (a city, castle), conquer (territory, a country); (c) to get (sth. to eat or drink); also, eat or drink (sth.); ~ **to drinken,** ~ **at eten**; (d) to obtain (a wind, a smell, hell, heaven, rain, a maidenhead, etc.); ~ **eft banke**, regain the bank; ~ **the feld**, be victorious; ~ the tour of London, be sent to the Tower.

7b. (a) To receive (a freed prisoner); rescue (a prisoner); bring (sb.) in for questioning; (b) to have (a woman) as wife or mistress; ~ **to (in, a) bed**, marry (a woman), have intercourse with (a woman); ~ **and holden**; also, to have (a man) as husband; (c) ~ **to (unto, as, for)**, to accept or receive (sb.) as (one's king, lord, superior); take (sb.) as (a witness, companion, wife); ~ **and holden**; -- also *refl.*; (d) to have (a child); give birth to (a child); beget (a child, an heir); ~ **upon**, beget (a child) on (a woman).

7c. To get (an abstraction); be the recipient of (an action): (a) to obtain (help, peace, mercy, favor, victory, etc.); receive (an answer); get (one's will), get (leave, permission); ~ **sight**, see (sb.), catch sight of (sth.); **houwe**, ?have we, ?may we have; -- also *refl.*; (b) to receive (harm, sorrow, punishment, a curse, etc.); ~ **deth**, die; ~ **vengaunce**, be punished by an avenger; ~ **while**, have a time of punishment; (c) to learn (sth.), find out, discover; comprehend (sth.), understand; (d) to receive (a blow, wound); have that, take that!; (e) to have (a beginning, an end, an origin); have (an ending of a certain kind); (f) in proverbs.

7d. In phrases, etc.: (a) ~ **the bettre (ende,** ~ **the heigher (over) hond,** ~ **the bataille**, to get the upper hand, be victorious, gain supremacy; ~ **the bettre of,** ~ **heigher hond of**, get the better of (sb.), overcome; ~ **the wors (ende,** ~ **the worst**, be overcome, be defeated; (b) ~ **god dai**, good-bye [see god dai]; (c) ~ **her mi feith (hond, trouth),** ~ **herof mi herte blod**, etc., I give you my word, etc.; (d) the devel (fend) have, the Devil take (sb. or sth.); Crist (God, our Lord) have, may Christ (God, our Lord) receive (someone's soul); (e) **what woldest thou** ~, what do you want?

8. (a) To pick up (sth.), take; (b) ~ **awei**, to take (sb. or sth.) away, drive away (an ache); ~ **doun**, knock or drag (sb.) to the ground; ~ **in**, bring in (hay), ?get in a stock of groceries; ~ **of**, take off (a helmet); ~ **out**, take (sb.) out of a place, remove (sth.), drive (sb.) out; ~ **up**, lift (sth.) up, hoist (a sail); ~ **forth,** ~ **her,** ~ **hom,** ~ **thider,** ~ **togeder**; (c) ~ **at**, to put (sth.) beside (sth.); ~ **awei from**, remove (sth.) from (sb.); ~ **fro**, take (sb.) from (sb.), remove (the head) from (the shoulders); ~ **in**, put (sb.) in (prison); ~ **into**, bring (sb.) into (a place); ~ **into place**, put (sth.) in place; ben had into, of a number: multiply (another number); ~ **of**, bring (sth.) from (a place); ~ **out of**, bring or carry (sb. or sth.) out of (a place); ben had out of place, be displaced; ~ **over**, put (sth.) over (sth.) as a cover; ~ **to**, bring or send (sb. or sth.) to (a person or place), invite (sb.) to (a meal, a feast); ~ **to ground**, topple (sth.) to the ground; ~ **unto,** ~ **up to**, bring (sb. or sth.) to (a place); (d) ~ **mid (with)**, to take (sb. or sth.) with (one); ~ **on**, take (sth.) on (a journey).

9. (a) To have or take (action); ~ **craft**, practice a trade; ~ **daunce**, dance; ~ **langage (speche)**, speak, talk; ~ **mong**, mingle; ~ **mot**, meet; ~ **plaie**, play; ~ **wordes**, speak words; etc.; (b) ~ **to don,** ~ **at do,** ~ **ado**, to have something to do, have business, be busy, be concerned; ~ **nought to don**, have nothing to do, have no business or concern; ~ **to don (ado) with**, have to do with (sth.), be concerned with, have dealings with (sb.); ~ **nought to don of (with),** ~ **not to don of,** ~ **nothing at do of**, have nothing to do with; (c) ~ **to don**, to do battle, fight; ~ **ado with**, fight with (sb.); (d) ~ **to don mid (with),** ~ **at do (ado) with**, to have intercourse with (a woman); (e) ~ **in commaundement**, to be commanded; ~ **in examinacioun**, question (sb.); ~ **in oth**, swear by (sth.); ~ **in sight**, see (sth.); ~ **in speche**, talk about (sth.); (f) in commands and exclamations: ~ **at**, to shoot at (sth.), strike at, attack; also, strike, attack, go to it; ~ **at the** (you), let me at you, here goes, let us fight; (g) ~ **don**, to finish; ?also, do away with (sb.); ~ **don with**, be finished with (sth.); in commands: finish; also, hurry, get on with it!

10. (a) To cause (sth.) to be (done), have (sth. done); have (sb. do sth.); -- usually with p.ppl.; -- also with inf. or clause; (b) to get (sb. or sth. into a state or condition); ~ **ded**, have (sb.) dead, kill; ~ **iler**, get (streets) cleared; (c) to cause (sb.) to be (somewhere, in a place or position, with sb.); -- also *refl.*; (d) to make (sth.), provide; have (sth.) made; (e) to allow (sth.), permit; (f) to suppose (sth.), assume.

11. (a) To have (sth. in order to do sth. with it); have (sb. or sth. to look after or attend to), have (sth. to say or write), have (a field to plow), etc.; (b) to be under obligation (to do sth.), have (to do sth.); ~ **to don,** ~ **to gon**, etc.; ~ **to don of (with)**, be under the necessity of dealing with (sb.); (c) to have occasion (to do sth.).

12a. Finite verb as auxiliary: (a) immediately after p.ppl.; (b) after p.ppl. & separated from it by subject; (c) immediately before p.ppl.; **yave graunted**, you have granted; (d) before p.ppl. & separated from it by subject (or subject & other elements of the clause); **havi**, I have; **hefdich**, had I, if I had; **haddich**, I would have; **nabbich**, I have not; **nefdich**, had I not; **hastou, havestu, hast thou, thou hast**; **haddestou, haddest thou**; (e) before p.ppl. & separated from it by an object or objects (or object & other elements of the clause); (f) before p.ppl. & separated from it by adverb(s or adv. phrases.

12b. Finite verb with ppl. **ben** or **haved**: (a) ~ **ben**, have existed, have been; -- often with neg.; (b) with adv., complement, or another ppl.; **hath ben thus**, has been thus; **han ben taken**, have been taken; **hath ben had**, has been had; etc.; (c) ~ **haved**, have had (sth.); -- in various senses.

12c. Infinitive as auxiliary or quasi-aux.: (a) **to-** or **forto-**inf. with p.ppl.; **to** ~ **ben hongen**, to be hanged; had **to have ben**, had been; **taben**, to have been; **taclipsed**, to have eclipsed; **tadaued** [see **adauen**], to have awakened; **tadwelled**, to have dwelt; **tafallen**, to have fallen; **talived**, to have lived; **tan touched**, to have touched; **tave holden**, to have held; (b) inf. without to: **had** ~ **ben**, had been; ~ **had ronnen**, to have run; (c) with modal & p.ppl.: **wolde** ~ **don, sholde** ~ **sen, couthe** ~ **boren, mighte** ~ **don**, etc.; (d) **mighte** ~ **ben, mighte** ~ **haved**; -- also with second p.ppl.

13. As a substitute for another verb, or with another verb (**ben, dien, don, dwellen, wenden**, etc.) implied.

C1 – Exemplification of the S01 for the 12th century (without observations).

TMN	GLOSS	FORMS OF *HAVEN*	FREQUENCY	TEXTS
H0	haven-INF	**habben**	**8**	**A030**
		habbe	2	A027
		habben	5	A027
		ha	1	A027
H0’	NEG-haven-INF	naffde	1	A030
H1	haven-PRS.1SG.IND	**habbe**	**18**	**A027 [7]; A030 [10]; A032 [1]**
		hafe	6	A028 [3]; A030 [3]
	haven-PRS.2SG.IND	hafesst	1	A030
	haven-PRS.3SG.IND	**haueþ**	**8**	**A009 [1]; A012 [1]; A013 [2]; A016 [1]; A018 [1]; A019 [1]; A027 [1]**
		hafeþþ	4	A028 [1]; A030 [3]
		haueð	6	A027
		habbe	1	A019
		efþ	1	A027
		hafð	1	A027
		hauet	1	A027
	haven-PRS.1PL.IND	habbeð	1	A027
		hefden	1	A027
		hafenn	2	A032
	haven-PRS.3PL.IND	hafenn	1	A030
		habbeð	**6**	**A027**
		habbeþ	1	A027
	NEG-haven-PRS.3SG.IND	nafð	1	A027
	NEG-haven-PRS.3PL.IND	nabbeð	2	A027
H2	haven-PST.2SG.IND	heuedest	1	A016
	haven-PST.3SG.IND	hadde	1	A007
		haffde	**24**	**A029 [5]; A030 [19]**
		had	2	A030
	haven-PST.3PL.IND	haffde	1	A030
		haffdenn	**13**	**A030 [12]; A032 [1]**
H2’	NEG-haven-PST.3SG.IND	naffde	2	A030
	NEG-haven-PST.3PL.IND	naffden	2	A030
H3	haven-PRS.2SG.SBJV	hafst	1	A012
		hauest	2	A014
		haue	1	A022
	haven-PRS.3SG.SBJV	haueþ	1	A007
		habben	1	A027
	haven-PRS.1PL.SBJV	habbe	1	A027
	haven-PRS.3PL.SBJV	habben	1	A025
H3’	NEG-haven-PRS.3SG.SBJV	nabbe	1	A022
H4	haven-PST.1SG.SBJV	hefde	1	A027
	haven-PST.3SG.SBJV	hefð	1	A027
		hefde	1	A027
	haven-PST.1PL.SBJV	hefden	1	A027
H4’	-	-	-	-
H5	-	-	-	-
H5’	-	-	-	-
H6	-	-	-	-
H7	-	-	-	-
Total			**138**	

C2 – Exemplification of the S01 for the 13th century (without observations).

TMN	GLOSS	FORMS OF *HAVEN*	FREQUENCY	TEXTS
H0	haven-INF	**have**	**9**	**B002 [2]; B007 [5]; B009 [2]**
		han	1	B003 [1]
		habbe	6	B067
		haven	3	B007 [2]; B022 [1]
		habben	2	B030
		hauen	1	B061
H0’	NEG-haven-INF	-	-	-
H1	haven-PRS.1SG.IND	**have**	**14**	**B007 [3]; B010 [5]; B014 [3]; B016 [3]**
		habbe	**13**	**B030 [1]; B067 [12]**
		haue	1	B060
	haven-PRS.2SG.IND	hast	4	B002 [1]; B009 [1]; B014 [1]
		haves	1	B007
		hauest	**13**	**B030 [3]; B067 [10]**
		haue	2	B030
		hauestu	1	B067
		habbe	1	B004
	haven-PRS.3SG.IND	hath	4	B002 [2]; B010 [2]
		haveth	5	B004 [4];B007 [1]
		haþ	1	B009 [1]
		haueð	1	B060
		habbe	5	B030 [1] B067 [2]
		hauet	1	B067
		haued	2	B067
		haueþ	**8**	**B067**
		hauen	1	B060
	haven-PRS.1PL.IND	hauen	1	B050
		haue	1	B050
		habbeþ	1	B067
	haven-PRS.3PL.IND	han	1	B003 [1]
		haveth	1	B004
		habbet	1	B067
		habbeþ	3	B067
		habeþ	1	B067
		haueþ	2	B067
		have	1	B071
H1’	NEG-haven-PRS.2SG.IND	neuestu	1	B067
		nauestu	1	B067
	NEG-haven-PRS.3SG.IND	naveth	2	B004
		nas	1	B009
		naueþ	5	B067
	NEG-haven-PRS.3PL.IND	nabbeþ	3	B067
		nabbed	1	B067
H2	haven-PST.1SG.IND	hadde	1	B067
	haven-PST.2SG.IND	haddest	1	B021
		hefedest	1	B030
	haven-PST.3SG.IND	havede	7	B007
		hadde	**16**	**B007 [1]; B009 [1]; B067 [14]**
		hedde	1	B030
	haven-PST.3PL.IND	hadde	2	B020 [1]; B067 [1]
H2’	NEG-haven-PST.2SG.IND	naddest	1	B067
	NEG-haven-PST.3SG.IND	nevede	1	B004

		nadde	2	B067
H3	haven-PRS.1SG.SBJV	habbe	3	B030
	haven-PRS.2SG.SBJV	hauest	2	B067
	haven-PRS.3SG.SBJV	hath	1	B020
		haueþ	1	B067
	haven-PRS.3PL.SBJV	habbeþ	1	B067
H3’	NEG-haven-PRS.2SG.SBJV	nauest	1	B030
H4	haven-PST.1SG.SBJV	hadde	1	B030
		heuede	1	B030
	haven-PST.2SG.SBJV	heuedest	1	B030
	haven-PST.3SG.SBJV	hadde	1	B067
	haven-PST.3PL.SBJV	hadde	1	B067
H4’	-	-	-	-
H5	haven-PRS.2SG.IMP	**have**	**5**	**B007 [2]; B009 [1]; B010 [1]; B021 [1]**
H5’	-	**-**	**-**	**-**
H6	-	**-**	**-**	**-**
H7	-	**-**	**-**	**-**
Total			**177**	

C3 – Exemplification of the S01 for the 14th century (without observations).

TMN	GLOSS	FORMS OF *HAVEN*	FREQUENCY	TEXTS
H0	haven-INF	**have**	**8**	**C014 [1]; C019 [1]; C020 [1]; C024 [1]; C033 [1]; C040 [2]; C050 [1]**
		haven	3	C016 [1]; C019 [1]; C020 [1]
		hafe	4	C024 [2]; C025 [2]
		han	1	C036
		ha	1	C043
		haf	1	C054
H0’	NEG-haven-INF	n’aven	1	C020
H1	haven-PRS.1SG.IND	**have**	**31**	**C003 [1]; C004 [3]; C005 [2]; C012 [2]; C013 [1]; C015 [3]; C016 [2]; C017 [1]; C027 [1]; C028 [1]; C040 [6]; C041 [1]; C043 [1]; C044 [1]; C046 [1]; C051 [1]; C054 [3]**
		Ichave	2	C010
		ha	2	C015
		hafe	1	C023
		haf	7	C054
		habbe	1	C054
	haven-PRS.2SG.IND	**hast**	**11**	**C001 [3]; C005 [2]; C015 [1]; C016 [1]; C027 [1]; C046 [2]; C051 [1]**
		hath	1	C003
		haste	2	C024
		hase	1	C024
		have	2	C024 [1]; C040 [1]
		hafe	2	C024 [1]; c025 [1]
		hastou	1	C027
		has	1	C028
		haf	2	C054
	haven-PRS.3SG.IND	**hath**	**36**	**C001 [1]; C003 [1]; C008 [1]; C009 [4]; C010 [1]; C012 [1]; C014 [4]; C015 [3]; C016 [6]; C018 [1]; C021 [4]; C034 [1]; C036 [3]; C040 [2]; C044 [1]; C047 [1]; C051 [1]**
		haveth	6	C007 [4]; C008 [1]; C016 [1]
		have	5	C011 [1]; C017 [1]; C018 [1]; C023 [1]; C039 [1]
		ha	1	C021
		hase	5	C023 [4]; C024 [1]
		has	6	C023 [1]; C024 [1]; C050 [4]
		hathe	2	C024 [1]; C040 [1];
		hafe	1	C025
		haven	1	C054
	haven-PRS.1PL.IND	haven	1	C020
	haven-PRS.2PL.IND	hafe	1	C024
		have	2	C039 [1]; C042 [1]
	haven-PRS.3PL.IND	**han**	**11**	**C001 [1]; C005 [1]; C006 [1]; C009 [2]; C020 [2]; C036 [1]; C051 [3]**
		hathe	1	C022

		hafe	2	C024 [1]; C025 [1]
		have	4	C027 [1]; C036 [1]; C050 [1]; C051 [1]
H1’	NEG-haven-PRS.1SG.IND	n’ave	2	C012 [1]; C041 [1]
	NEG-haven-PRS.3SG.IND	nath	3	C020 [1]; C021 [1]; C051 [1]
H2	haven-PST.1SG.IND	hade	2	C012 [1]; C020 [1]
		hadde	1	C043
	haven-PST.3SG.IND	hadde	2	C002 [1]; C046 [1]
		had	**6**	**C008 [1]; C023 [2]; C050 [2]; C054 [1]**
		hade	**5**	**C020 [1]; C054 [4]**
		hevez	1	C054
	haven-PST.2PL.IND	haden	1	C024
	haven-PST.3PL.IND	had	2	C025
H2’	NEG-haven-PST.2PL.IND	nade	1	C025
H3	haven-PRS.1SG.SBJV	ha	1	C015
	haven-PRS.2SG.SBJV	have	1	C012
		haste	1	C024
		haf	1	C054
	haven-PRS.3SG.SBJV	**hath**	**6**	**C011 [3]; C020 [1]; C036 [1]; C051 [1]**
	haven-PRS.3PL.SBJV	hath	1	C024
H3’	-	-	-	-
H4	haven-PST.1SG.SBJV	hadde	1	C043
	haven-PST.3SG.SBJV	hadde	2	C014 [1]; C051 [1]
		hade	1	C053
		had	2	C053 [1]; C054 [1]
	haven-PST.3PL.SBJV	haden	1	C054
		hade	1	C054
H4’	NEG-haven-PST.3SG.SBJV	nade	1	C053
H5	haven-PRS.2SG.IMP	**have**	**5**	**C002 [1]; C003 [1]; C005 [1]; C017 [1]; C034 [1]**
		ha	1	C008
	haven-PRS.2PL.IMP	have	1	C042
H6	haven-PRS.PTCP	havende	1	C024
H7	haven-PST.PTCP	y-had	1	C018
Total			**226**	

C4 – Exemplification of the S01 for the 15th century (without observations).

TMN	GLOSS	FORMS OF *HAVEN*	FREQUENCY	TEXTS
H0	haven-INF	**haue**	**33**	**D003 [1]; D006 [4]; D007 [1]; D010 [2]; D011 [1]; D017 [1]; D02D003 [1]; D006 [4]; D007 [1]; D010 [2]; D011 [1]; D017 [1]; D020 [1]; D022 [1]; D024 [1]; D025 [1]; D031 [4]; D032 [4]; D033 [1]; D034 [2]; D035 [3]; D036 [2]; D037 [1]; D038 [2]0 [1]; D022 [1]; D024 [1]; D025 [1]; D031 [4]; D032 [4]; D033 [1]; D034 [2]; D035 [3]; D036 [2]; D037 [1]; D038 [2]**
		have	1	D006
		hawe	1	D035
H0’	NEG-haven-INF	-	-	-
H1	haven-PRS.1SG.IND	have	3	D001 [2]; D002 [1]
		haue	**13**	**D007 [1]; D008 [1]; D025 [4]; D031 [2]; D034 [2]; D035 [2]; D036 [1]**
		haff	1	D029
		hes	1	D033
	haven-PRS.2SG.IND	**hast**	**8**	**D005 [1]; D007 [1]; D016 [1]; D019 [1]; D025 [1]; D034 [3]**
		has	2	D007 [1]; D035 [1]
		hase	2	D007 [1]; D017 [1]
		haste	2	D011 [1]; D035 [1]
		haue	3	D025 [1; D026 [1]; D034 [1]
		hes	1	D034
	haven-PRS.3SG.IND	**hath**	**14**	**D002 [1]; D006 [1]; D017 [1]; D020 [2]; D025 [1]; D031 [5]; D036 [1]; D037 [2]**
		as	1	D004
		has	**11**	**D004 [1]; D006 [2]; D007 [2]; D012 [1]; D021 [3]; D035 [2]**
		haþe	1	D011
		hes	1	D014
		hase	2	D017 [1]; D019 [1]
		haþ	2	D025 [1]; D038 [1]
		hasse	1	D029
		haues	1	D013
	haven-PRS.1PL.IND	haue	2	D032 [1]; D038 [1]
	haven-PRS.2PL.IND	hase	1	D007
		haue	1	D010
	haven-PRS.3PL.IND	has	2	D007 [1]; D024 [1]
		haue	**5**	**D020 [1]; D023 [1]; D025 [1] D031 [2]**
		hath	1	D032
		han	1	D036
H1’	-	-	-	-
H2	haven-PST.1SG.IND	hadde	2	D006 [1]; D022 [1]
		had	2	D006 [1]; D025 [1]
		hade	3	D030

	haven-PST.2SG.IND	hadest	2	D011
		hadyst	1	D020
	haven-PST.3SG.IND	hedde	1	D006
		had	**9**	**D010 [1]; D012 [3]; D024 [1]; D029 [1]; D033 [2]; D034 [1]**
		hat	1	D036
	haven-PST.2PL.IND	had	2	D010
	haven-PST.3PL.IND	had	3	D029 [1]; D030 [1]; D033 [1]
		hat	1	D036
H2’	NEG-haven-PST.3SG.IND	nad	1	D024
H3	haven-PRS.1SG.SBJV	haue	2	D031 [1]; D025 [1]
	haven-PRS.2SG.SBJV	hes	2	D033
	haven-PRS.3SG.SBJV	haue	2	D003 [1]; D031 [1]
		haþ	2	D038
	haven-PRS.3PL.SBJV	han	1	D038
H3’	-	-	-	-
H4	haven-PST.1SG.SBJV	had	4	D031 [2]; D035 [1]; D038 [1]
	haven-PST.2SG.SBJV	hadde	1	D034
	haven-PST.3SG.SBJV	had	3	D007 [1]; D034 [1]; D035 [1]
		hat	1	D036
	haven-PST.2PL.SBJV	hade	1	D010
	haven-PST.3PL.SBJV	had	2	D006
H4’	-	-	-	-
H5	haven-PRS.2SG.IMP	**haue**	**10**	**D006 [1]; D008 [1]; D013 [1]; D026 [1]; D027 [4]; D033 [1]; D037 [1]**
H5’	-	**-**	**-**	**-**
H6	-	**-**	**-**	**-**
H7	haven-PST.PTCP	**had**	**5**	**D025 [1]; D031 [1]; D032 [1]; D033 [1]; D035 [1]**
Total			**183**	

D1 – Representation of the S02 for the 12th century.

Text code	Forms of *haven*	Occurrence number	Gloss	Classification
A005	habbe	#1	Uncertain	Uncertain
A007	hadde	#1	haven-PST.3SG.IND	Grammatical
	haueþ	#2	haven-PRS.3SG.SBJV	Uncertain
A009	haueþ	#1	haven-PRS.3SG.IND	Grammatical
A012	hafst	#1	haven-PRS.2SG.SBJV	Lexical
	haueþ	#2	haven-PRS.3SG.IND	Lexical
A013	haueþ	#1	haven-PRS.3SG.IND	Uncertain
	haueþ	#2	haven-PRS.3SG.IND	Uncertain
A014	hauest	#1	haven-PRS.2SG.SBJV	Lexical
	hauest	#2	haven-PRS.2SG.SBJV	Lexical
A016	heuedest	#1	haven-PST.2SG.IND	Grammatical
	haueþ	#2	haven-PRS.3SG.IND	Lexical
A018	haueþ	#1	haven-PRS.3SG.IND	Grammatical
A019	habbe	#1	haven-PRS.3SG.IND	Lexical
	haueþ	#2	haven-PRS.3SG.IND	Lexical
A022	haue	#1	haven-PRS.2SG.SBJV	Lexical
	nabbe	#2	NEG-haven-PRS.3SG.SBJV	Uncertain

D2 – Representation of the S02 for the 13th century.

Text code	Forms of *haven*	Occurrence number	Gloss	Classification
B002	have	#1	haven-INF	Lexical
	hast	#2	haven-PRS.2SG.IND	Lexical
	hath	#3	haven-PRS.3SG.IND	Grammatical
	have	#4	haven-INF	Lexical
	hast	#5	haven-PRS.2SG.IND	Grammatical
	hath	#6	haven-PRS.3SG.IND	Grammatical
B003	han	#1	haven-PRS.3PL.IND	Lexical
	han	#2	haven-INF	Lexical
B004	haveth	#1	haven-PRS.3SG.IND	Lexical
	haveth	#2	haven-PRS.3PL.IND	Grammatical
	nevede	#3	NEG-haven-PST.3SG.IND	Lexical
	naveth	#4	NEG-haven-PRS.3SG.IND	Lexical
	naveth	#5	NEG-haven-PRS.3SG.IND	Lexical
	haveth	#6	haven-PRS.3SG.IND	Grammatical
	habbe	#7	haven-PRS.2SG.IND	Grammatical
	haveth	#8	haven-PRS.3SG.IND	Grammatical
	haveth	#9	haven-PRS.3SG.IND	Grammatical

D3 – Representation of the S02 for the 14th century.

Text code	Forms of *haven*	Occurrence number	Gloss	Classification
C001	hast	#1	haven-PRS.2SG.IND	Grammatical
	hast	#2	haven-PRS.2SG.IND	Grammatical
	han	#3	haven-PRS.3PL.IND	Lexical
	hath	#4	haven-PRS.3SG.IND	Grammatical
	hast	#5	haven-PRS.2SG.IND	Grammatical
C002	hadde	#1	haven-PST.3SG.IND	Lexical
	have	#2	haven-PRS.2SG.IMP	Lexical
C003	hath	#1	haven-PRS.3SG.IND	Grammatical
	hath	#2	haven-PRS.2SG.IND	Grammatical
	have	#3	haven-PRS.2SG.IMP	Lexical
	have	#4	haven-PRS.1SG.IND	Grammatical
C004	have	#1	haven-PRS.1SG.IND	Grammatical
	have	#2	haven-PRS.1SG.IND	Grammatical
	have	#3	haven-PRS.1SG.IND	Grammatical

D4 – Representation of the S02 for the 15th century.

Text code	Forms of *haven*	Occurrence number	Gloss	Classification
D001	have	#1	haven-PRS.1SG.IND	Grammatical
	have	#2	haven-PRS.1SG.IND	Lexical
D002	have	#1	haven-PRS.1SG.IND	Grammatical
	hath	#2	haven-PRS.3SG.IND	Grammatical
D003	haue	#1	haven-INF	Lexical
	haue	#2	haven-PRS.3SG.SBJV	Grammatical
D004	as	#1	haven-PRS.3SG.IND	Grammatical
	has	#2	haven-PRS.3SG.IND	Grammatical
D005	hast	#1	haven-PRS.2SG.IND	Lexical
D006	hadde	#1	haven-PST.1SG.IND	Grammatical
	had	#2	haven-PST.1SG.IND	Grammatical
	hedde	#3	haven-PST.3SG.IND	Grammatical
	haue	#4	haven-INF	Grammatical
	haue	#5	haven-INF	Grammatical
	has	#6	haven-PRS.3SG.IND	Lexical
	haue	#7	haven-INF	Grammatical
	has	#8	haven-PRS.3SG.IND	Lexical
	had	#9	haven-PST.3PL.SBJV	Grammatical
	haue	#10	haven-INF	Grammatical
	had	#11	haven-PST.3PL.SBJV	Grammatical
	hath	#12	haven-PRS.3SG.IND	Grammatical
	haue	#13	haven-PRS.2SG.IMP	Lexical
	have	#14	haven-INF	Lexical

E1 – Script for lexical occurrences sampling – 13th century.

```
# Sampling lexical occurrences - 13th century

############# Preliminary steps
# show objects in the memory
ls()
# clean memory
rm(list=ls())
# check current directory
getwd()
# show files from directory
dir()
#############

# read file from the 13th century
file1 = read.csv2("sampling_13th_century.csv")
file1

# filter lexical occurrences
file1_lex = subset (file1, Classification == "Lexical")
file1_lex

# sample
# sample size
sample_size = 55

# population size
population_size = nrow(file1_lex)
population_size

# sampling
set.seed(666)
sample1 = sample(seq_along(1:population_size),sample_size)
sample1

# sort sample
sample1 = sort (sample1)
sample1

# sampled occurrences
samples_13_cent = file1_lex [sample1,]
samples_13_cent

# number of occurrences
nrow(samples_13_cent)

# export results
write.csv2(samples_13_cent,"samples_13_cent.csv",row.names=F)

#####################################################
```

E2 – Script for lexical occurrences sampling – 14th century.

```
# Sampling lexical occurrences - 14th century

############# Preliminary steps
# show objects in the memory
ls()
# clean memory
rm(list=ls())
# check current directory
getwd()
# show files from directory
dir()
#############

# read file from the 14th century
file1 = read.csv2("sampling_14th_century.csv")
file1

# filter lexical occurrences
file1_lex = subset (file1, Classification == "Lexical")
file1_lex

# sample
# sample size
sample_size = 53

# population size
population_size = nrow(file1_lex)
population_size

# sampling
set.seed(666)
sample1 = sample(seq_along(1:population_size),sample_size)
sample1

# sort sample
sample1 = sort (sample1)
sample1

# sampled occurrences
samples_14_cent = file1_lex [sample1,]
samples_14_cent

# number of occurrences
nrow(samples_14_cent)

# export results
write.csv2(samples_14_cent,"samples_14_cent.csv",row.names=F)

#####################################################
```

E3 – Script for lexical occurrences sampling – 15th century.

```
# Sampling lexical occurrences - 15th century

############# Preliminary steps
# show objects in the memory
ls()
# clean memory
rm(list=ls())
# check current directory
getwd()
# show files from directory
dir()
#############

# read file from the 15th century
file1 = read.csv2("sampling_15th_century.csv")
file1

# filter lexical occurrences
file1_lex = subset (file1, Classification == "Lexical")
file1_lex

# sample
# sample size
sample_size = 53

# population size
population_size = nrow(file1_lex)
population_size

# sampling
set.seed(666)
sample1 = sample(seq_along(1:population_size),sample_size)
sample1

# sort sample
sample1 = sort (sample1)
sample1

# sampled occurrences
samples_15_cent = file1_lex [sample1,]
samples_15_cent

# number of occurrences
nrow(samples_15_cent)

# export results
write.csv2(samples_15_cent,"samples_15_cent.csv",row.names=F)

#####################################################
```

F1-A – 12th century's readings with their respective diatheses (Part I).

haven	*verb*		
1a(a)	**'to possess'**		
	Active		
[N]	Arg1 V Arg2		
	Arg2 Arg1 V		
	Arg1 Arg2 V		
	Arg1 Arg1' V Arg2		(in relative clauses)
1c	**'to possess (sth. together with the capacity to do sth. with it)'**		
	Active		
[N to-INF]	Arg2 Arg1 V to-INF		
2(b)	**'in fig. phrases: be able to speak'**		
	Active		
[N ADJ]	Arg1 + V Arg2 ADJ	coll. ~ tunge	
[N]	Arg1 V Arg2		
2(d)	**'to contain (sth.)'**		
	Active		
[N]	Arg1 Arg1' V Arg2		(in relative clauses)
3(a)	**'to have (sb.) under one)'**		
	Active		
[N]	V Arg1 Arg2		
3(c)	**'to have (sb.) in a certain relationship'**		
	Active		
[N]	Arg 2 Arg 2 Arg1 V		(in relative clauses)
[N PP]	Arg1 V Arg2 Arg3		
3(d)	**'to have (a father, mother, son, daughter, husband, wife, etc.)'**		
	Active		
[N]	Arg1 V Arg2		
4a(e)	**'to have (a law, commandment; (...) have an agreement with (sb.)**		
	Active		
[N PP]	Arg2 Arg1 Arg3 V		
4b(a)	**'to have (sth. that denotes a state or condition)'**		
	Active		
[N]	Arg1 V Arg2		
	Arg2 Arg1 V		
	V Arg1 Arg2		
[N PP]	Arg3 Arg1 V Arg2		
4b(f)	**'be in trouble'**		
	Active		
[N]	Arg2 V Arg1		
4b(g)	**'be barely able (to do sth.)'**		
	Active		
[V]	V(haven) Arg1 V(PTCP)	This V is a particular case.	

F1-B – 12th century's readings with their respective diatheses (Part II).

4c(a)	**'to have (a nature, quality, characteristic, etc.)'**		
	Active		
[N]	Arg1 V Arg2		
[N for-to-INF]	V Arg1 Arg2 for-to-INF		
4c(d)	**'to have a moral or spiritual quality'**		
	Active		
[N]	Arg1 Arg1' V Arg2		
	V Arg1 Arg2		
4c(e)	**'show mercy to (sb. or sth.), take pity on'**		
	Active		
[N]	Arg1 V Arg2		
4d(a)	**'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'**		
	Active		
[N]	Arg1 V Arg2		
[N PP]	Arg1 V Arg2 Arg3		
	Arg1 Arg3 V Arg2		
[N INF]	Arg1 V Arg2 INF		
4d(c)	**'have interest in (sth.)'**		
	Active		
[N PP]	Arg1 V Arg2 Arg3		
4d(e)	**'~ in wille, intend (to do sth.)'**		
	Active		
[PP INF]	Arg1 V Arg3 INF		
7b(d)	**'to have (a child); give birth to (a child); beget (a child, an heir)'**		
	Active		
[N]	V Arg1 Arg2		
	Arg1 Arg2 V		
7c(a)	**'to obtain (sth. abstract)'**		
	Active		
[N]	V Arg1 Arg2		
	Arg1 Arg2 V		
	Arg1 Arg1' V Arg2		(in relative clauses)
	Arg1 V Arg2		
[N PP]	Arg1 Arg3 V Arg2		
	Arg1 Arg1' Arg2 V Arg3		(in relative clauses)
[N to-INF(haven)]	Arg1 Arg1' V(to-INF(haven)) Arg2		(in relative clauses)
[N for-INF(haven)]	Arg1 Arg1' V(for-INF(haven)) Arg2		(in relative clauses)
7c(b)	**'to receive (harm, sorrow, punishment, a curse, etc.)'**		
	Active		
[N]	Arg1 V Arg2		
11(b)	**'to be under obligation (to do sth.)'**		
	Active		
[to-INF]	V Arg1 INF		

G - List 1 - Examples of the 12th century's readings.

1a(a) – 'to possess'

[N] – *ech mon mid he* ***hauet*** *mei buggen houene riche*
each man with he has may buy heaven kingdom

"each man is able to buy the kingdom of heaven with all he has" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

1c – 'to possess (sth. together with the capacity to do sth. with it)'

[N to-INF] – *wel muchel he hauet to beten*
well much he has to repent

"he has much to repent" - Arg2 Arg1 V to-INF

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with the verbal particle 'to' (to-INF).

2(b) – 'in fig. phrases: be able to speak'

[N ADJ] – *wymmon is word.woþ and haueþ tunge to swift*
woman is word-crazy and has tongue too swift

"woman is word-crazy, and has a swift tongue" - Arg1 V Arg2 ADJ

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Adjective (ADJ) to modify the Direct Object for this figurative sentence to make the desired sense.

[N] – *many mon haueþ swikelne muþ*
many a man has treacherous mouth

"many a man has a treacherous mouth" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

2(d) – 'contain (sth.)'

[N] – *quaþþrigan tthat hafeþþ fowwre wheless*
chariot that has four wheels

"a chariot that has four wheels" - Arg1 Arg1' V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2). As there is a relative clause, Arg1 is duplicated as Arg1'.

3(a) – 'to have (sb.) under one'

[N] – *and hafesst ȝet elldermanness láte*
and you_have yet royal_officers a_lot

"and yet you have a lot of royal officers" - V Arg1 Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

3(c) – 'to have (sb.) in a certain relationship'

[N] – *freond þat he habbe*
friend that he has

"a friend that he has" - Arg2 Arg2' Arg1 V

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2). As there is a relative clause, Arg2 is dueplicated as Arg2'.

[N PP] – *aȝȝ þeȝȝ haffdenn allderrmenn and kingess off hem sellfenn*
always they had rulers and kings over themselves

"they always had rulers and kings over themselves" - Arg1 V Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and a reflexive Indirect Object (Arg3).

3(d) – 'to have (a father, mother, son, daughter, husband, wife, etc.)'

[N] – *Zacariȝe haffde an duhhtiȝ wif*
Zechariah had a virtuous wife

"Zechariah had a virtuous wife" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

4a(e) – 'to have (a law, commandment); (…) have an agreement with (sb.)'

[N PP] – *ne mot he of þære molde habben nammore*
no discussion they about that land have no_more

"if they have no more discussion about that land" - Arg2 Arg1 Arg3 V

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Oblique Object (Arg3). In addition, this interpretation and translation only approximates the subsense given by MED coded as *4a(e)*.

4b(a) – 'to have (sth. that denotes a state or condition)'

[N] – *nabbeð hi nane blisse*
not_have they no peace

"[they both (cold and heat) do them enough suffering;] they have no peace" - V Arg1 Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

[N PP] – *of hete hi habbeð misse*
of heat they have need

"[when they come again to cold,] they have need of heat" - Arg3 Arg1 V Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Oblique Object (Arg3).

4b(f) – 'be in trouble'
[N] – *monies monnes sare iswinc habbeð oft unholde*
many a_man suffering trouble have often unfriendly

"for the unfriendly (ones) often have many a man's suffering trouble" - Arg2 V Arg1

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

4b(g) – 'be barely able (to do sth.)'
[V] – *naffde ȝho nohht tæmedd*
not_had she not begotten

"for she was a woman so consisted that she had not to beget" - V(haven) Arg1 V(PTCP)

Observation: This reading requires a verb in the past participle form as well as a Subject (Arg1). Despite its structure, *haven*'s function is not as a grammatical verb, requiring a participle to form past tense. The choice here was to keep the meaning, not the metric or original structure, what justifies the translation above. This is why this case is special and thus marked in yellow in Appendix F1-A.

4c(a) – 'to have (a nature, quality, characteristic, etc.)'
[N] – *forr Crist iss baþe godd and mann an had off twinne kinde*
for Christ is both god and man and has of two natures

"For Christ is both man and god and has both natures" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

[N for-to-INF] – *þatt naffde ȝho nan kinde þa onn hire for to tæmenn*
that not_had she no nature the in her for to beget

"so that she had no nature in her then for generating" - V Arg1 Arg2 for-to-INF

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with the verbal particles 'for' and 'to' (for-to-INF).

4c(d) – 'to have a moral or spiritual quality'
[N] – *forr alle þa þatt hafenn witt*
for all the that have sense

"for all those who have sense" - Arg1 Arg1' V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2). As there is a relative clause, Arg1 is duplicated as Arg1'.

4c(e) – ‘show mercy to (sb. or sth.), take pity on’
[N] – *sulf douel mihte habben milce ȝif he hit bigunne*
self devil could have mercy if he it begun

"the devil himself could have mercy if he (had) begun it" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

4d(a) – ‘to experience (a feeling, fear, pleasure, wonder, joy, etc.)’
[N] – *ach þo þre habbeþ scome and grome and oft fele sorȝe*
but the others have shame and wrath and other great sorrows

"But others have shame and wrath and other great sorrows" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

[N PP] – *whatt menn mihhtenn habenn niþ ne wraþþe ȝæn heore owwþerr*
what men could have envy neither wrath against each other

"what men might have neither envy nor wrath against each other" - Arg1 V Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Indirect Object (Arg3).

[N INF] – *alle shulenn habbenn blisse tosammenn stanndenn*
all shall have joy together to_stay

"everyone shall have joy to stay together" - Arg1 V Arg2 INF

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with no particle such as ‘to’ or the combination ‘for to’ (INF).

4d(c) – ‘have interest in (sth.)’
[N PP] – *þer ne þert he habben kare of ȝeve ne of ȝelde*
there no to_need he have interest in reward no in gold

"it is not necessary that he has concern of reward or tribute" - Arg1 V Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and a compound Oblique Object (Arg3).

4d(e) – ‘~ in wille, intend (to do sth.)’
[PP INF] – *þatt ȝho nohht naffde inn hire wille to cnawenn aniȝ macchess stren*
that she not not_had in her consciousness to know any men’s offspring

"[if that might be so,] not have in her consciousness to know any man's offspring" - Arg1 V Arg3 INF

Observation: This reading requires a Subject (Arg1), an Oblique Object (Arg3) and an infinitive clause functioning as a Direct Object.

7b(d) – 'to have (a child); give birth to (a child); beget (a child, an heir)'

[N] – *þatt naffdenn þeȝȝ nan child till þa*
that not_have they no child until then

"(You can see ...) that they had no child until then" - V Arg1 Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

7c(a) – 'to obtain (sth. abstract)'

[N] – *þenne haueð he his mid iwisse*
then has he it with certainty

"let him do good himself while he may, then he will have it [God's bliss] with certainty" - V Arg1 Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

[N PP] – *in hure blisse þe þe ende haueð for endelese pine*
in our bliss which the end has for endless pain

"And our bliss, which has an end for endless pain" - Arg1 Arg1' Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Oblique Object (Arg3). As there is a relative clause, Arg1 is duplicated as Arg1'.

[N to-INF(haven)] – *þe mon þe wule siker bon to habben goddes blisse*
the man who wishes certain to_be to have God's bliss

"the man who wishes to be certain to have God's bliss" - Arg1 Arg1' V(to-INF(haven)) Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with the verb *haven* and the particle 'to' (to-INF(haven)). As there is a relative clause, Arg1 is duplicated as Arg1'.

[N for-INF(haven)] – *þo þe er doð eni god for habben godes are*
he who here does any good to have God's mercy

"he who does any good here to have God's mercy" - Arg1 Arg1' V(to-INF(haven)) Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with the verb *haven* and the particle 'for' (for-INF(haven)). As there is a relative clause, Arg1 is duplicated as Arg1'.

7c(b) – 'to receive (harm, sorrow, punishment, a curse, etc.)'

[N] – *þe sculen habbe herdne dom her weren herde*
those shall have hard judgement here were hard

"Those who were hard here shall have a hard judgement" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

11(b) – ‘to be under obligation (to do sth.)’

[to-INF] – *ne haue þu to vale worde*
no have you to remove word

"if you have to remove (a) word" - V Arg1 INF

Observation: This reading requires a Subject (Arg1) and an infinitive clause with the verbal particle ‘to’ (to-INF).

H - List 2. Readings of *haven* for the 12th century according to the MED ordered into the most to the least frequent ones considering the sampled lexical occurrences.

7c(a) - 'to obtain (sth. abstract)' - 10 occ.
1a(a) - 'to possess' - 8 occ.
4d(a) - 'to experience (a feeling, fear, pleasure, wonder, joy, etc.)' - 5 occ.
3(c) - 'to have (sb.) in a certain relationship' - 4 occ.
4b(a) - 'to have (sth. that denotes a state or condition)' - 4 occ.
4c(a) - 'to have (a nature, quality, characteristic, etc.)' - 3 occ.
2(b) - 'in fig. phrases: be able to speak' - 2 occ. ok
7c(b) - 'to receive (harm, sorrow, punishment, a curse, etc.)' - 2 occ.
4c(d) - 'to have (a moral or spiritual quality)' - 2 occ.
7b(d) - 'to have (a child); give birth to (a child); beget (a child, an heir)' - 2 occ.

HAPAX LEGOMENA

1c - 'to possess (sth. together with the capacity to do sth. with it)'
2(d) - 'contain (sth.)'
3(a) - 'to have (sb.) under one'
3(d) - 'to have (a father, mother, son, daughter, husband, wife, etc.)'
4a(e) - 'to have (a law, commandment); be bound by (an oath); have (an agreement with sb.)'
4b(f) - 'be in trouble'
4b(g) - 'be barely able (to do sth.)'
4c(e) - 'show mercy to (sb. or sth.), take pity on'
4d(c) - 'have interest in (sth.)'
4d(e) - '~ in wille, intend (to do sth.)'
11(b) - 'to be under obligation (to do sth.)'

I1-A – 13th century's readings with their respective diatheses (Part I).

haven	*verb*		
1a(a)	**'to possess'**		
	Active		
[N]	Arg2 Arg1 V		
	Arg1 V Arg2		
[N INF]	Arg2 Arg2' Arg1 V(INF-haven)		(in relative clauses)
1a(e)	**'to have (sth.) in (one's) possession or under (one's) control'**		
	Active		
[N PP]	Arg1 V Arg3 Arg2		
1b(a)	**'to have (sb. or sth. somewhere)'**		
	Active		
[N]	Arg1 Arg1' V Arg2		(in relative clauses)
[N PP]	Arg1 V Arg2 Arg3		
1c	**'to possess (sth. together with the capacity to do sth. with it)'**		
	Active		
[N to-INF]	Arg2 Arg1 V Arg2 to-INF		
2(a)	**'to be provided with (a part, an organ)'**		
	Active		
[N]	Arg1 V Arg2		
	Arg1 Arg1' V Arg2		(in relative clauses)
2(d)	**'contain (sth.)'**		
	Active		
[N PP]	Arg1 V Arg2 Arg3		
3(d)	**'to have (a father, mother, son, daughter, husband, wife, etc.)'**		
	Active		
[N]	Arg1 V Arg2		
3(e)	**'~ evening (felaue, per), to have an equal'**		
	Active		
[]	no arguments	coll. ~ euening	
[ALC]	no arguments	coll. ~ evenyng	
4b(a)	**'to have (sth. that denotes a state or condition)'**		
	Active		
[N]	Arg1 V Arg2		
	Arg1 Arg2 V		
	Arg2 V Arg1		
[N to-INF(haven)]	Arg1 V(to-INF(haven)) Arg2		
[N PP]	Arg1 V Arg2 Arg3		
4b(b)	**'suffer (pain, harm, a blow, a wound)'**		
	Active		
[N PP]	Arg1 V Arg2 Arg3	fig. construction	
4b(d)	**'~ nede of (to), to need (sth.)'**		
	Active		
[N PP]	Arg3 V Arg1 Arg2	coll. ~ ned	

I1-B – 13th century's readings with their respective diatheses (Part II).

4c(b)	**'to be endowed with (a physical quality)'**		
	Active		
[N]	Arg1 V Arg2		
4c(c)	**'to have (a function, power, etc.)'**		
	Active		
[N]	V-Arg1 Arg2		
	Arg1 V Arg2		
4c(d)	**'to have (a moral or spiritual quality)'**		
	Active		
[N]	Arg1 V Arg2		
[N INF(haven)]	Arg1 V(INF-haven) Arg2		
4c(e)	**'show mercy to (sb. or sth.), take pity on'**		
	Active		
[N PP]	V-Arg1 Arg2 Arg3	coll. ~ merci of	
4d(a)	**'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'**		
	Active		
[N]	Arg1 V Arg2		
[N PP]	Arg3 V-Arg1 Arg2		
6a(a)	**'to regard (sb. or sth. in a certain way), consider'**		
	Active		
[N CL]	V Arg1 Arg2 CL		
7a(a)	**'to obtain (sth.)'**		
	Active		
[N]	V-Arg1 Arg2		
7a(c)	**'to get (sth. to eat or drink)'**		
	Active		
[N]	Arg1 V Arg2		
7c(a)	**'to obtain (sth. abstract)'**		
	Active		
[N]	Arg1 V Arg2		
	Arg2 V Arg1		
[N INF(haven)]	Arg1 V(INF-haven) Arg2		
7c(b)	**'to receive (harm, sorrow, punishment, a curse, etc.)'**		
	Active		
[CL]	Arg1 V CL		
[N]	Arg1 V Arg2		
9(b)	**'have dealings with (sb.)'**		
	Active		
[N PP]	Arg1 Arg3 V Arg2		
	NOT CONTEMPLATED BY MED - TYPE 1: 'to be quiet'		
	Active		
[N PP]	V Arg1 Arg2 Arg3		
	NOT CONTEMPLATED BY MED - TYPE 2: 'to have an answer'		
	Active		
[N]	Arg1 V Arg2		

J – List 3. Examples of the 13th century's readings.

1a(a) – 'to possess'

[N] – *hevy hameres they han*
heavy hammers they have

"they have heavy hammers" - Arg2 Arg1 V

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

[N INF] – *the plates that he haven wolde*
the plates that he to_have wished

"the plates that he wished to have" - Arg2 Arg2' Arg1 V

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2). As there is a relative clause, Arg2 is duplicated as Arg2'.

1a(e) – 'to have (sth.) in (one's) possession or under (one's) control'

[N PP] – *he shal haven in his handel al Denemark and Engeland*
he shall have in his hands al Denmark and England

"he shall have both Denmark and England under his power" - Arg1 V Arg3 Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Oblique Object (Arg3).

1b(a) – 'to have (s. or sth. somewhere)'

[N] – *suche wede that naveth king ne kayser non*
such wedding that had_not king not emperor not

"[He brought you to] such wedding that had neither king nor emperor" - Arg1 Arg1' V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2). As there is a relative clause, Arg1 is duplicated as Arg1'.

[N PP] – *men habbet a rum-hus at hore bures ende*
men have a private_place in their house's end

"men have a private place hidden in their houses" - Arg1 V Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Oblique Object (Arg3).

1c – 'to possess (sth. together with the capacity to do sth. with it)'

[N to-INF] – *þin armes he haþ and scheld to fiȝte wiþ*
thin arms he has and shield to fight with

"he has thin arms and shield to fight with" - Arg2 Arg2 V to-INF

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with the verbal particle 'to' (to-INF).

2(a) – ‘to be provided with (a part, an organ)’

[N] – *þu hauest wel scharpe clawe*
you have well sharp claws

"you have well sharp claws" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

2(d) – ‘contain (sth.)’

[N PP] – *Ich habbe at wude tron wel grete*
I have at wood tree well big

"I have wood in a big tree" - Arg1 V Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Oblique Object (Arg3).

3(d) – ‘to have (a father, mother, son, daughter, wife, etc.)’

[N] – *þar louerd haueþ his loue ibedde*
where lord has his beloved spouse

"where the lord has his beloved spouse" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

3(e) – ‘~ evening (felaue, per), to have an equal’

[] – *ac wisdom naueþ non euening*
-- wisdom not_have no equal

"there was no equal wisdom" - no arguments

Observation: This reading requires no argument since this meaning is existential.

[ALC] – *Henry ure king that nevede on eorthe non evenyng*
Henry our king that had_not in earth no equal

"Henry, our king, that there was no equal on this earth" - ALC

Observation: This reading requires an Allocutive (ALC) and this meaning is existential.

4b(a) – ‘to have (sth. that denotes a state or condition)’

[N] – *Havelok mihte sei þat him ne havede grip or ern*
Havelok could say that him not had skill or honour

"[Havelok could say] that he had neither skill nor honour" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

[N to-INF(haven)] – *me is lof to habbe reste*
me is important to have rest

"for me it is important to have rest" – Arg1 V(to-INF-haven) Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with the verb *haven* and the verbal particle ‘to’ (to-INF-haven).

4b(b) – 'to suffer (pain, harm, a blow, a wound)'
[N PP] – *heo haueþ þe fust in hire teþ*
she has the fist in her teeth

"she has his fist in her teeth" - Arg1 V Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Oblique Object (Arg3). Besides, this is a figurative construction.

4b(d) – '~ need of (to), to need (sth.)'
[N PP] – *ure louerdes lage her of haue we mikel ned*
our Lord's law here of have we much need

"we have much need of our Lord's law" - Arg1 V Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Oblique Object (Arg3).

4c(b) – 'to be endowed with (a physical quality)'
[N] – *ich habbe gode sene*
I have good sight

"[It is said about me that] I have a good sight" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

4c(c) – 'to have (a function, power, etc.)'
[N] – *think man whilstou hast mighte*
think man while_you have power

"think, man, while you still have power" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

4c(d) – 'to have (a moral or spiritual quality)'
[N] – *and þar men habbeþ milde mod*
and where men have kind heart

"and where men have kind heart" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

[N INF(haven)] – *and nu ich ho pie habben fulle uorȝiuenesse*
and now I cease merciful have full forgiveness

"And now I cease to have merciful, full forgiveness" - Arg1 V(INF-haven) Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with the verb *haven* (INF-haven).

4c(e) – 'show mercy to (sb. or sth.), take pity on'
[N PP] – *Loverd have merci of me*
Lord have mercy on me

"Lord, have mercy on me!" - V-Arg1 Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Indirect Object (Arg3). Besides, this construction has 'merci of' as a collocate of *haven*.

4d(a) – 'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'
[N] – *he havede michel shame*
he had much shame

"he had much shame" – Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

[N PP] – *luve God almighty and of him have drede*
love God Almighty and of him have fear

"love the God Almighty and have fear of Him" - Arg3 V-Arg1 Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Indirect Object (Arg3).

6a(a) – 'to regard (sb. or sth. in a certain way)'
[N CL] – *naueþ no man siker-hede þat he ne mai wene*
not_have no man certainty that he not may distress

"no man has any certainty that he may not distress" - V Arg1 Arg2 CL

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and a *that* Clause (CL).

7a(a) – 'to obtain (sth.)'
[N] – *have her þis cuppe*
have here this goblet

"have here this goblet" - V-Arg1 Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

7a(c) – 'to get (sth. to eat or drink)'
[N] – *shaltu have non oþer mede*
you_shall have no other food

"You shall have no other food" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

7c(a) – 'to obtain (sth. abstract)'

[N] – *ich habbe boþe luue and þonc*
I have both love and thanks

"I have both love and thanks" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

[N INF(haven)] – *ich bidde þat men beon iwarre an habbe gode reads*
I beg that men be aware and have good advice

"I beg men to be vigiland and have good advice" - Arg1 V(INF-haven) Arg2

This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with the verb *haven* (INF-haven).

7c(b) – 'to receive (harm, sorrow, punishment, a curse, etc)'

[CL] – *thou shalt **have** as thou hast wrought*
you shall **have** as you have done

"you shall **have** as you have done" - Arg1 V CL

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an *as* Clause (CL).

[N] – *þu naddest non oþer dom ne laȝe*
you not_have no other judgement no sentence

"you had neither judgement nor sentence" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

9(b) – 'have dealings with (sb.)'

[N PP] – *Alured seide þat wit þe fule haueþ i-mene*
Alfred is_said that with the filthy has association

"Alfred is said to have association with filthy a person" - Arg1 Arg3 V Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Indirect Object (Arg3).

NOT CONTEMPLATED BY MED – TYPE 1 – 'to be quiet'

[N PP] – *habbe he is tunge under gore*
have he his tongue under clothing

"[For when he has done his deed] he shall have his tongue under clothing" - V Arg1 Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Oblique Object (Arg3). Besides, this is a figurative construction.

NOT CONTEMPLATED BY MED – TYPE 2 – 'to have an answer'

[N] – *an ȝet ich habbe an oþer andsware*
and yet I have another answer

"and yet I have another answer" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

K – List 4. Readings of *haven* for the 13th century according to the MED ordered into the most to the least frequent ones considering the sampled lexical occurrences.

4b(a) - 'to have (sth. that denotes a state or condition)' - 9 occ.
2(a) - 'to be provided with (a part, an organ)' - 5 occ.
1a(a) - 'to possess' - 4 occ.
4c(b) - 'to be endowed with (a physical quality)' - 4 occ.
4d(a) - 'to experience (a feeling, fear, pleasure, wonder, joy, etc.)' - 4 occ.
7c(a) - 'to obtain (sth. abstract)' - 4 occ.
1b(a) - 'to have (sb. or sth. somewhere)' - 2 occ.
3(e) - '~ evening (felaue, per), to have an equal' - 2 occ.
4c(c) - 'to have (a function, power, etc.)' - 2 occ.
4c(d) - 'to have (a moral or spiritual quality)' - 2 occ.
NOT CONTEMPLATED BY MED - TYPE 2 – 'to have an answer': 2 occ.

HAPAX LEGOMENA

1a(e) - 'to have (sth.) in (one's) possession or under (one's) control'
1c - 'to possess (sth. together with the capacity to do sth. with it)'
2(d) - 'contain (sth.)'
3(d) - 'to have (a father, mother, son, daughter, husband, wife, etc.)'
4b(b) - 'suffer (pain, harm, a blow, a wound)'
4b(d) - '~ nede of (to), to need (sth.)'
4c(e) - 'show mercy to (sb. or sth.), take pity on'
6a(a) - 'to regard (sb. or sth. in a certain way), consider'
7a(a) - 'to obtain (sth.)'
7a(c) - 'to get (sth. to eat or drink)'
7c(b) - 'to receive (harm, sorrow, punishment, a curse, etc.)'
9(b) - 'have dealings with (sb.)'
NOT CONTEMPLATED BY MED - TYPE 1 – 'to be quiet'

L1-A – 14th century's readings with their respective diatheses (Part I).

haven	***verb***		
1a(a)	**'to possess'**		
	Active		
[N]	Arg1 V Arg2		
	Arg2 Arg2' Arg1 V		(in relative clauses)
	Arg2 Arg1 V		
	Arg1 Arg1' V Arg2		(in relative clauses)
1a(e)	**'to have (sth.) in (one's) possession or under (one's) control'**		
	Active		
[N PP CL]	Arg1 V Arg2 Arg3 CL		
1c	**'to possess (sth. together with the capacity to do sth. with it)'**		
	Active		
[N to-INF]	Arg1 Arg1' V Arg2 to-INF		(in relative clauses)
[N for-to-INF [PP]]	Arg1 V Arg2 for-to-INF		
2(a)	**'to be provided with (a part, an organ)'**		
	Active		
[N]	Arg2 Arg1 V		
	Arg1 V Arg2		
2(c)	**'~ on (upon), have on (garment, etc.)'**		
	Active		
[N PP for-to-INF]	Arg2 Arg1 V on for-to-INF	coll. ~ on	
2(d)	**'contain (sth.)'**		
	Active		
[N]	Arg1 V Arg2		
3(a)	**'to have (sb.) under one'**		
	Active		
[N]	Arg1 Arg1' V Arg2		(in relative clauses)
	Arg2 Arg1 V		
3(b)	**'have power over (sb.)'**		
	Active		
[N PP]	Arg3 Arg1 V Arg2		
3(c)	**'to have (sb.) in a certain relationship'**		
	Active		
[N PP]	Arg1 Arg1' V Arg3 Arg2		(in relative clauses)
3(e)	**'~ evening (felaue, per), to have an equal'**		
	Active		
[ALC]	no arguments	coll. ~ pere	
4a(a)	**'to enjoy (a right or privilege), have (the right to possess sth.); have (power, etc.); have (leave or license to do sth.)'**		
	Active		
[N]	Arg1 V Arg2		
4a(h)	**'~ part (parti) of, have anything to do with (sb. or sth.)'**		
	Active		
[N PP]	Arg1 V Arg2 Arg3	coll. ~ parte	
4b(a)	**'to have (sth. that denotes a state or condition)'**		
	Active		
[N]	Arg1 V Arg2		
4c(c)	**'to have (a function, power, etc.), (~ sight, be able to see; ~ lif, have life, be alive)'**		
	Active		
[N INF(haven) PP]	Arg1 Arg3 V(INF-haven) Arg2	coll. ~ sight	
[N]	Arg1 V Arg2	coll. ~ lyves	

L1-B – 14th century's readings with their respective diatheses (Part II).

4c(d)	**'to have (a moral or spiritual quality)'**		
	Active		
[N]	Arg1 V Arg2	coll. ~ sin	
4c(e)	**'show mercy to (sb. or sth.), take pity on'**		
	Active		
[N PP]	V-Arg1 Arg2 Arg3	coll. ~ mercy of, merci of	
	Arg1 Arg1' V Arg3 Arg2	coll. ~ on ... mercy	(in relative clauses)
	Arg1 Arg1' Arg3 V Arg2	coll. ~ of... pite	(in relative clauses)
4d(a)	**'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'**		
	Active		
[N]	Arg1 V Arg2		
	Arg2 Arg1 V		(Arg1 repeated twice after V)
	V Arg1 Arg2		
[N PP]	Arg1 Arg2 V Arg3		
	Arg3 V-Arg1 Arg2		
	Arg3 V Arg1 Arg2		
4d(c)	**~ minde, to remember (sth.), consider**		
	Active		
[N PP]	V-Arg1 Arg2 Arg3	coll. ~ mynde	
4d(e)	**'~ in herte(wit), to feel (sth.) in the heart (mind)'**		
	Active		
[PP CL]	Arg3 Arg1 V CL	coll. ~ in ... hertes	
5b(a)	**'to keep (sb. or sth. in a place)'**		
	Active		
[N PP]	Arg1 V Arg3 Arg2		
6a(a)	**'to regard (sb. or sth. in a certain way), consider'**		
	Active		
[N PP]	Arg1 Arg3 V Arg2		
7a(a)	**'to obtain (sth.)'**		
	Active		
[to-INF(haven) PP]	Arg1 V(to-INF(haven)) Arg3		
[N]	Arg1 V(INF-haven) Arg2		
7a(d)	**'to obtain (a wind, a smell, hell, heaven, rain, a maidenhead, etc.)'**		
	Active		
[N PP [CL]]	Arg1 V Arg2 P [CL]		
7c(a)	**'to obtain (sth. abstract)'**		
	Active		
[N to-INF(haven)]	Arg1 Arg2 to-INF(haven)		
[N]	Arg1 V Arg2		
7c(b)	**'to receive (harm, sorrow, punishment, a curse, etc.)'**		
	Active		
[N]	Arg2 V Arg1 Arg1'		(in relative clauses)
7c(d)	**'to receive (a blow, wound)'**		
	Active		
[N ADJ]	ADJ Arg1 V Arg2	fig. construction	
9(b)	**'~ nought to don, have nothing to do, have no business or concern'**		
	Active		
[N to-INF PP]	Arg1 Arg3 V Arg2 to-INF	coll. ~ nought to don	
10(b)	**'to get (sb. or sth. into a state or condition)'**		
	Active		
[N ADJ]	Arg1 V Arg2 ADJ		
11(b)	**'to be under obligation (to do sth.)'**		
	Active		
[N to-INF]	Arg1 Arg2 V to-INF		
[N]	Arg1 V Arg2		

M – List 5. Examples of the 14th century's readings.

1a(a) – 'to possess'

[N] – *thenne mot Ich have hennes arost*
then can I have hens roast

"then I can have roast hens" - Arg2 Arg1 V

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

1a(e) – 'to have (sth.) in (one's) possession or under (one's) control'

[N PP CL] – *I haf hit holly in my honde þat al desyres þurȝe grace*
I have it wholly in my hand what all you_desire Your Grace

"I have it wholly under my control what you desire, Your Grace" - Arg1 V Arg2 Arg3 CL

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2), an Oblique Object (Arg3) and a *that* Clause (CL).

1c – 'to possess (sth. together with the capacity to do sth. with it)'

[N to-INF] – *Hilde that haveth me to hede*
Hilde who has me to care_for

"[More gracious that] Hilde, who has me to care for" - Arg1 Arg1' V Arg2 to-INF

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with the verbal particle 'to' (to-INF). As there is a relative clause, Arg1 is duplicated as Arg1'.

[N for-to-INF [PP]] – *I have nother clout ne cloth thee inne for to folde*
I have neither rag nor cloth you in for to wrap

"I have neither rag nor cloth to dress you up" - Arg1 V Arg2 for-to-INF

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with the combination of the verbal particles 'for' and 'to' (for-to-INF).

2(a) – 'to be provided with (a part, an organ)'

[N] – *middle she hath menskful smal*
waist she has beautiful slender

"she has a beautiful, slender waist" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

2(c) – '~ on (upon), have on (garment, etc.)'

[N PP for-to-INF] – *clothes y have on for to caste*
clothes I have on for to put

"I have clothes to put on" - Arg2 Arg1 V on for-to-INF

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2), the collocate preposition 'on' and an infinitive clause with the combination of the verbal particles 'for' and 'to' (for-to-INF).

2(d) – 'contain (sth.)'

[N] – *the breste had another [brod chechun]*
the chest had another [large shield-shaped_ornament]

"[With a well-fitting tunic fastened at the sides, a large shield-shaped ornament in the back]; the chest had another" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

3(a) – 'to have (sb.) under one'

[N] – *and iche a segge that I see has sexe mens doke*
and each visitor that I see has six men's amount

"and each visitor I see has the amount of six men" - Arg1 Arg1' V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2). As there is a relative clause, Arg1 is duplicated as Arg1'.

3(b) – 'have power over (sb.)'

[N PP] – *of maidnes meke thou hast mighte*
over maidens submissive you have power

"[Richard, source of good sense ...] you have control over submissive maidens" - Arg3 Arg1 V Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Indirect Object (Arg3).

3(c) – 'to have (sb.) in a certain relationship'

[N PP] – *he that have to fere his moste fo*
he who has as wife his worst enemy

"he who has as a wife his worst enemy" - Arg1 Arg1' V Arg3 Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Indirect Object (Arg3). As there is a relative clause, Arg1 is duplicated as Arg1'.

3(e) – '~ evening (felaue, per), to have an equal'

[ALC] – *the sonne bryght that of yelownesse hadde never pere*
the sun bright that of yellowness had never equal

"the bright sun that had no equal yellowness" - ALC

Observation: This reading requires an Allocutive (ALC) and this meaning is existential. Besides, this construction has 'pere' as a collocate of *haven*.

4a(a) – 'to enjoy (a right or privilege), have (the right to possess sth.); have (power, etc.); have (leave or license to do sth.)'

[N] – *this may haveth might*
this maiden has power

"this maiden has power" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

4a(h) – '~ part (parti) of, have anything to do with (sb. or sth.)'

[N PP] – *late the peple and the pore hafe parte of thi silvere*
let the people and the poor have part of your silvery

"let the people and the poor ones have part of your silvery" - Arg1 V Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Oblique Object (Arg3). Besides, this is a figurative construction.

4b(a) – 'to have (sth. that denotes a state or condition)'

[N] – *sithe that I have this sorowe withoute hir rede*
since that I have this sorrow without her encouragement

"since I have this sorrow without her encouragement, [then may I say shortly that her womanhood is not to blame]" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

4c(c) – 'to have (a function, power, etc.)'

[N INF(haven) PP] – *God us lene His light that we of seyntes haven sight*
God us grant His light that we of saints to_have sight

"may God grant us His light for us to have sight of the saints" - Arg1 Arg3 V(INF-haven) Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Indirect Object (Arg3). Besides, this construction has 'sight' as a collocate of *haven*.

[N] – *for wel I wot whyle ye have lyves space*
for well I know while you have lifespan

"I know, while you have your lifespan" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2). Besides, this construction has 'lyves' as a collocate of *haven*.

4c(d) – 'to have (a moral or spiritual quality)'

[N] – *she haveth sin*
she has sin

"[Unless she loves me] she has sin" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2). Besides, this construction has 'sin' as a collocate of *haven*.

4c(e) – 'show mercy to (sb. or sth.), take pity on'
[N PP] – *lady ha mercy of thy man*
lady have mercy on your man

"Lady, have mercy on your man" - V-Arg1 Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Indirect Object (Arg3). Besides, this construction has 'mercy of' as a collocate of *haven*.

4d(a) – 'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'
[N] – *I have no wonder though yee do me woo*
I have no astonishment though you do me grief

"I have no astonishment though you cause me grief" – Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

[N PP] – *to me hastou non awe*
for me you_have no fear

"for me you have no fear" – Arg3 V-Arg1 Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Indirect Object (Arg3).

4d(c) – '~ minde, to remember (sth), consider'
[N PP] – *have mynde upon my supplicacion*
have mind upon my supplication

"have my supplication in mind" - V-Arg1 Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Oblique Object (Arg3). Besides, this construction has 'mynde' as a collocate of *haven*.

4d(e) – '~ in herte (wit), to feel (sth.) in the heart (mind)'
[PP CL] – *but in her hertes I wolde they hade hou sone that Gode hem may degrade*
but in their hearts I wished they had how straightaway that God them may bring_down

"but I wish they had it in their hearts how straightaway that God may bring them down" – Arg3 Arg1 V CL

Observation: This reading requires a Subject (Arg1), an Oblique Object (Arg3) and a *how* Clause (CL). Besides, this construction has 'in … hertes' as a collocate of *haven*.

5b(a) – 'to keep (sb. or sth. in a place)'
[N PP] – *and syþen I have in þis hous hym þat al lykez*
and since I have in this house him that all pleases

"and since I have in this house all that pleases him" - Arg1 V Arg3 Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Indirect Object (Arg3).

6a(a) – 'to regard (sb. or sth. in a certain way)'
[N PP] – *though men to me han onde*
though men towards me have enmity

"though men have enmity towards me" - Arg1 Arg3 V Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Oblique Object (Arg3).

7a(a) – 'to obtain (sth.)'
[to-INF(haven) PP] – *the hayward heteth us harm to have of his*
the hayward causes us trouble to have of his

"the hayward causes us trouble to have his bit" - Arg1 V Arg3 Arg2 to-INF(haven)

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2), an Indirect Object (Arg3) and an infinitive clause with the verb *haven* and the verbal particle 'to' (to-INF(haven)).

[N] – *þe stablye þat let þe herttez haf þe gate*
the militia that let the harts have the gate

"[The militia] that let the harts have the gate" - Arg1 V(INF-haven) Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

7a(d) – 'to obtain (a wind, a smell, hell, heaven, rain, a maidenhead, etc.)'
[N PP[CL]] – *and thou hafe helle full hotte for that thou here saved*
and you have hell quite hot for what you here saved

"and you have a quite hot hell for what you saved here" – Arg1 V Arg2 P [CL]

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and the preposition 'for' followed by a *that* Clause (CL).

7c(a) – 'to obtain (sth. abstract)'
[N to-INF(haven)] – *God yeve thee grace god happes to have*
God give you grace good fortune to have

"[Sir,] may God give you grace to have good fortune" - Arg1 Arg2 V(to-INF-haven)

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with the verb *haven* and the verbal particle 'to' (to-INF(haven)).

[N] – *he had his will at Berwik*
he had his will at Berwik

"he had his will at Berwik" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

7c(b) – 'to receive (harm, sorrow, punishment, a curse, etc)'

[N] – *hethyng have the hathell that any harme thynkes*
shame have the warrior who any harm thinks

"may the warrior who thinks badly of it have shame" - Arg2 V Arg1 Arg1'

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2). As there is a relative clause, Arg1 is duplicated as Arg1'.

7c(d) – 'to receive (a blow, wound)'

[N ADJ] – *open thou hast thi syde*
open you have your side

"you have your side open" - ADJ Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2). Besides, this is a figurative construction whose context tells that Jesus' side is open due to his wound.

9(b) – 'have dealings with (sb.)'

[N to-INF PP] – *for she of my folye hath nought to done although she do me starve*
for she of my foolishness has not to do although she causes me to_die

"for she has nothing to do with my foolishness although she causes me to die" - Arg1 Arg3 V Arg2 to-INF

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2), an Oblique Object (Arg3) and an infinitive clause with the verbal particle 'to' (to-INF). Besides, this construction has 'nought to done' as a collocate of *haven*.

10(b) – 'to get (sb. or sth. into a state or condition)'

[N ADJ] – *weping hath myn wonges wet*
crying has my cheeks wet

"crying has my cheeks wet" - Arg1 V Arg2 ADJ

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Adjective (ADJ).

11(b) – 'to be under obligation (to do sth.)'

[N to-INF] – *sithe I counte and cot hade to kepe*
since I accounts and small_holding had to keep

"Since I had accounts and a small holding to keep" - Arg1 Arg2 V to-INF

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with the verbal particle 'to' (to-INF).

[N] – *and al o deth bos bothe drye and han on incarnacioun*
and all one death be_obliged both suffer and have one birth

"and all must both suffer one death and have one birth" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

N – List 6. Readings of *haven* for the 14th century according to the MED ordered into the most to the least frequent ones considering the sampled lexical occurrences.

1a(a) - 'to possess' - 7 occ.
4d(a) - 'to experience (a feeling, fear, pleasure, wonder, joy, etc.)' - 6 occ.
4c(e) - 'show mercy to (sb. or sth.), take pity on' 4 occ.
2(a) - 'to be provided with (a part, an organ)' - 3 occ.
7c(a) - 'to obtain (sth. abstract)' - 3 occ.
1c - 'to possess (sth. together with the capacity to do sth. with it)' - 2 occ.
3(a) - 'to have (sb.) under one' - 2 occ.
4c(c) - 'to have (a function, power, etc.) (~ sight, be able to see; ~ lif, have life, be alive)' - 2 occ.
4d(c) - 'to remember (sth.), consider' (1st. ~ minde) - 2 occ.
7a(a) - 'to obtain (sth.)' - 2 occ.
11(b) - 'to be under obligation (to do sth.)' - 2 occ.

HAPAX LEGOMENA

1a(e) - 'to have (sth.) in (one's) possession or under (one's) control'
2(c) - '~ on (upon), have on (a garment, etc.)'
2(d) - 'contain (sth.)'
3(b) - 'have power over (sb.)'
3(c) - 'to have (sb.) in a certain relationship'
3(e) - '~ evening (felaue, per), to have an equal'
4a(a) - 'to enjoy (a right or privilege), have (the right to possess sth.); have (power etc); have (leave or license to do sth.)'
4a(h) - '~ part (parti) of, have anything to do with (sb. or sth.)'
4b(a) - 'to have (sth. that denotes a state or condition)'
4c(d) - 'to have (a moral or spiritual quality)'
4d(e) - '~ in herte (wit), to feel (sth.) in the heart (mind)'
5b(a) - 'to keep (sb. or sth. in a place)'
6a(a) - 'to regard (sb. or sth. in a certain way), consider'
7a(d) - 'to obtain (a wind, a smell, hell, heaven, rain, a maidenhead, etc.)'
7c(b) - 'to receive (harm, sorrow, punishment, a curse, etc.)'
7c(d) - 'to receive (a blow, wound)'
9(b) - '~ nought to don, have nothing to do, have no business or concern'
10(b) - 'to get (sb. or sth. into a state or condition)'

O1-A – 15th century's readings with their respective diatheses (Part I).

haven	***verb***		
1a(a)	**'to possess'**		
	Active		
[N]	Arg1 Arg1' V Arg2		(in relative clauses)
1b(a)	**'to have (sb. or sth. somewhere)'**		
	Active		
[N PP]	Arg1 V Arg2 Arg3		
	Arg2 Arg3 Arg1 V		
1c	**'to possess (sth. together with the capacity to do sth. with it)'**		
	Active		
[N to-INF]	Arg1 V Arg2 to-INF		
1d(a)	**'to have (sb. or sth. in a state or condition)'**		
	Active		
[N PP]	Arg1 V Arg2 Arg3		
2(a)	**'to be provided with (a part, an organ)'**		
	Active		
[N]	Arg2 Arg1 V to-INF		
3(e)	**'~ evening (felaue, per), to have an equal'**		
	Active		
[]	Arg1 V	coll. ~ pere	
4a(a)	**'to enjoy (a right or privilege), have (the right to possess sth.); have (power, etc.); have (leave or license to do sth.)'**		
	Active		
[N to-INF]	Arg1 V Arg2 to-INF		
4a(g)	**'~ place, of a medicine: have a proper time to be used'**		
	Active		
[N to-INF]	Arg1 V Arg2 to-INF	coll. ~ place	
4b(a)	**'to have (sth. that denotes a state or condition)'**		
	Active		
[N]	Arg1 V Arg2		
[N PP]	Arg3 Arg1 V Arg2		
4b(a) & 4b(b)	**'to have (a disease)' & 'to have (hostility)'**		
	Active		
[N]	Arg2 Arg2' Arg1 V		(in relative clauses)
4b(b)	**'suffer (pain, harm, a blow, a wound)'**		
	Active		
[N PP]	Arg1 V Arg2 Arg3		
4b(c)	**'to have (cause or reason), have (reason to do sth.)'**		
	Active		
[N to-INF]	Arg2 V Arg1 to-INF		
	Arg1 V Arg2 to-INF		
4c(a)	**'to have (a nature, quality, characteristic, etc.)'**		
	Active		
[N]	Arg2 V Arg1		
4c(c)	**'to have (a function, power, etc.), (~ sight, be able to see; ~ lif, have life, be alive)'**		
	Active		
[N to-INF(haven) PP]	Arg3 V(to-INF(haven)) Arg2	coll. ~ syȝte	
[N]	Arg1 V Arg2	coll. ~ lyff	

O1-B – 15th century's readings with their respective diatheses (Part II).

haven	***verb***		
1a(a)	**'to possess'**		
	Active		
[N]	Arg1 Arg1' V Arg2		(in relative clauses)
1b(a)	**'to have (sb. or sth. somewhere)'**		
	Active		
[N PP]	Arg1 V Arg2 Arg3		
	Arg2 Arg3 Arg1 V		
1c	**'to possess (sth. together with the capacity to do sth. with it)'**		
	Active		
[N to-INF]	Arg1 V Arg2 to-INF		
1d(a)	**'to have (sb. or sth. in a state or condition)'**		
	Active		
[N PP]	Arg1 V Arg2 Arg3		
2(a)	**'to be provided with (a part, an organ)'**		
	Active		
[N]	Arg2 Arg1 V to-INF		
3(e)	**'~ evening (felaue, per), to have an equal'**		
	Active		
[]	Arg1 V	coll. ~ pere	
4a(a)	**'to enjoy (a right or privilege), have (the right to possess sth.); have (power, etc.); have (leave or license to do sth.)'**		
	Active		
[N to-INF]	Arg1 V Arg2 to-INF		
4a(g)	**'~ place, of a medicine: have a proper time to be used'**		
	Active		
[N to-INF]	Arg1 V Arg2 to-INF	coll. ~ place	
4b(a)	**'to have (sth. that denotes a state or condition)'**		
	Active		
[N]	Arg1 V Arg2		
[N PP]	Arg3 Arg1 V Arg2		
4b(a) & 4b(b)	**'to have (a disease)' & 'to have (hostility)'**		
	Active		
[N]	Arg2 Arg2' Arg1 V		(in relative clauses)
4b(b)	**'suffer (pain, harm, a blow, a wound)'**		
	Active		
[N PP]	Arg1 V Arg2 Arg3		
4b(c)	**'to have (cause or reason), have (reason to do sth.)'**		
	Active		
[N to-INF]	Arg2 V Arg1 to-INF		
	Arg1 V Arg2 to-INF		
4c(a)	**'to have (a nature, quality, characteristic, etc.)'**		
	Active		
[N]	Arg2 V Arg1		
4c(c)	**'to have (a function, power, etc.), (~ sight, be able to see; ~ lif, have life, be alive)'**		
	Active		
[N to-INF(haven) PP]	Arg3 V(to-INF(haven)) Arg2	coll. ~ syȝte	
[N]	Arg1 V Arg2	coll. ~ lyff	

P – List 7. Examples of the 15th century's readings.

1a(a) – 'to possess'
[N] – *y am lafte here as a woman forsake þat no goode has*
I am left here as a woman forsaken who no wealth has

"I am left here as a forsaken woman that has no wealth" - Arg1 Arg1' Arg2 V

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

1b(a) – 'to have (sb. or sth. somewhere)'
[N PP] – *when ȝe play and hase your childur on kness daunsand*
when you play and have your children on knees bouncing

"[Behold, women,] when you play and have your children bouncing on your knees" - Arg1 V Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Oblique Object (Arg3).

1c – 'to possess (sth. together with the capacity to do sth. with it)'
[N to-INF] – *þer he hath pour to reyse and rowte*
there it has power to raise and bellow

"it has power to raise and bellow" - Arg1 V Arg2 to-INF

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with the verbal particle 'to' (to-INF).

1d(a) – 'to have (sb. or sth. in a state or condition)'
[N PP] – *whanne age haþ us at his auauntage*
when age has us at its advantage

"then we may not do a big deal, [but sometimes groan, and sometimes complain, and sometimes scratch itching pustules] when age has us at its advantage" - Arg1 V Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Oblique Object (Arg3).

2(a) – 'to be provided with (a part, an organ)'
[N] – *and also erys þu hast to here*
and also ears you have to hear

"and also ears you have to hear" - Arg2 Arg1 V

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

3(e) – '~ evening (felaue, per), to have an equal'

[] – *if ȝe luf youres [son] myne has no pere*
if you love your [son] mine has no equal

"if you love your son, (then) mine has no equal" - Arg1 V

Observation: This reading requires a Subject (Arg1) and no more arguments since this meaning is existential. Besides, this construction has 'pere' as a collocate of *haven*.

4a(a) – 'to enjoy (a right or privilege), have (the right to possess sth.); have (power, etc.); have (leave or license to do sth.)'

[N to-INF] – *shall we neuer haue lycence to lyve yn ese*
shall we never have licence to live in peace

"we shall never have licence to live in peace" - Arg1 V Arg2 to-INF

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with the verbal particle 'to' (to-INF).

4a(g) – '~ place; of a medicine: have a proper time to be used'

[N to-INF] – *I haue no place to represse þem aright*
I have no time to repress them properly

"I have no time to repress them [the wounds] properly" - Arg1 V Arg2 to-INF

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with the verbal particle 'to' (to-INF). Besides, this construction has 'place' as a collocate of *haven*.

4b(a) – 'to have (sth. that denotes a state or condition)'

[N] – *in þat hyþe place þu hast hovnowre*
in that heavenly place you have honour

"in that heavenly place you have honour" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

[N PP] – *seynt george on whom all englond hath byleve*
Saint George on whom all England has faith

"Saint George, on whom all England has faith, [show us your help to the Almighty God, and keep our king from all mischief]" - Arg3 Arg1 V Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Indirect Object (Arg3).

4b(a) & 4b(b) – 'to have (a disease)' & 'to have (hostility)'

[N] – *and al oþer dissese or stryff þat I haue* ***had*** *seþ I had first lyff*
and all other diseases or hostility that I have **had** since I had first life

"and all other illnesses or hostility that I have **had** since I had first life" - Arg2 Arg2' Arg1 V

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2). As there is a relative clause, Arg2 is duplicated as Arg2'.

4b(b) – 'suffer (pain, harm, a blow, a wound)'

[N PP] – *myn enmys saide I hade the devyll at my demayne*
my enemies said I had the devil at my control

"my enemies said I had the devil inside me" - Arg1 V Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Oblique Object (Arg3).

4b(c) – 'to have (cause or reason); have (reason to do sth.)'

[N to-INF] – *what cause hast þu so sore to wepe*
what cause have you so sadly to cry

"what reason do you have to cry so sadly?" - Arg2 V Arg1 to-INF

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with the verbal particle 'to' (to-INF).

4c(a) – 'to have (a nature quality, characteristic, etc.)'

[N] – *me herytage modur haue þu shall*
my heritage mother have you shall

"My heritage, mother, you shall have it" - Arg2 V Arg1

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

4c(c) – 'to have (a function, power, etc.)'

[N INF(haven) PP] – *hertely y beseche of my swete sone to have a syȝte*
sincerely I command of my sweet son to have a sight

"I command all my guidance to have a sight of my sweet son" - Arg3 V(to-INF(haven)) Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Indirect Object (Arg3). Besides, this construction has 'syȝte' as a collocate of *haven* meaning 'to see'.

[N] – *and al oþer dissese or stryff þat I haue had seþ I **had** first lyff*
and all other diseases or hostility that I have had since I **had** first life

"and all other ilnesses or hostility that I have had since I **had** first life" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2). Besides, this construction has 'lyff' as a collocate of *haven* meaning 'to be alive'.

4c(d) – 'to have (a moral or spiritual quality)'

[N] – *o lady þat arte so bright as is þe sunne ffor þe gret ioye þou hadest in sight*
o lady who is so bright as is the sun for the great joy you had in sight

"o lady, you are so bright as is the sun for the great joy you had in sight" - Arg2 Arg1 V

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

[N to-INF(haven)] – *ffor gold makus many a man in gode or euel to haue no tast*
for gold makes many a man in good or evil to have no sensibility

"for gold turns many a man into good or evil to have no sensibility" - Arg1 V to-INF(haven) Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with the verb *haven* and the verbal particle 'to' (to-INF(haven)).

4c(e) – 'show mercy to (sb. or sth.), take pity on'
[N PP] – *ffader one thi sonne haue pety*
father on your son have pity

"father, have pity on your son" - Arg3 V-Arg1 Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Indirect Object (Arg3). Besides, this construction has 'one ... pety' as a collocate of *haven*.

[N] – *Cayam mycht haue had marcy weill aneuch*
Cain could have had mercy well enough

"Cain could have had mercy well enough" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

4d(a) – 'to experience (a feeling, fear, pleasure, wonder, joy, etc.)'
[N] – *ȝe thoȝt ȝe hade a full gode game*
you thought you had a very good pleasure

"you thought you had a good pleasure" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

[N CL] – *a secunde ioy hadyst þu whan cryst ihu of þe was borne*
a second joy had you when Jesus Christ of you was born

"a second joy you had when Jesus Christ was born of you" - Arg2 V Arg1 CL

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and a *when* Clause (CL).

4d(b) – 'to have (an idea, a thought), have (knowledge, etc.)'
[N] – *if þu be a lytill chyld ȝitt may þu haue þi wyll*
if you are a little child yet may you have your consciousness

"if you are a little child you may yet have your (own) consciousness" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

4d(c) – '~ minde, to remember (sth), consider'

[N to-INF(haven) PP] – *and I prey þi sone to haue my sowle in his lovely eie*
and I prey your son to have my soul in his lovely eye

"and I pray your son to have my soul in his lovely eye" - Arg1 V(to-INF(haven)) Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2), an Oblique Object (Arg3) and an infinitive clause with the verb *haven* and the verbal particle 'to' (to-INF(haven)).

[N PP to-INF] – *ffor of man he had a goode remembraunce to suffre deth*
for of mankind he had a good remembrance to suffer death

"for he had a good remembrance of people to suffer death" - Arg3 Arg1 V Arg2 to-INF

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2), an Indirect Object (Arg3) and an infinitive clause with the verbal particle 'to' (to-INF). Besides, this construction has 'remembraunce' as a collocate of *haven*.

[N PP] – *but yt one þis þu hast no mynde*
but it on this you have no mind

"but you do not remember it" - Arg3 Arg1 V Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Oblique Object (Arg3). Besides, this construction has 'one ... mynde' as a collocate of *haven*.

6a(a) – 'to regard (sb. or sth. in a certain way)'

[N PP ADV] – *but þair askyng in alkin thing thai had als fast*
but their request in everything they had as certain

"but they had their request for everything just as certain" - Arg2 Arg3 Arg1 V ADV

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2), an Oblique Object (Arg3) and the adverbial construction 'als fast' (ADV).

6a(c) – '~ in hate, hate (sb.)'

[N PP] – *oure frendys þat schul loue vs beste þan wol haue vs but in hate*
our friends who shall love us dearly then want to_have us but in hate

"our friends who should love us dearly will then have us but in hate" - Arg1 Arg1' V(INF(haven) Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2), an Indirect Object (Arg3) and an infinitive clause with the verb *haven* (INF(haven)). Besides, this construction has 'in hate' as a collocate of *haven*. As there is a relative clause, Arg1 is duplicated as Arg1'.

7c(a) – 'to obtain (sth. abstract)'

[N PP] – *thou woful moder and mayde hadest deþe in þy dolour*
you afflicted mother and maiden you_had death in your suffering

"o you, mother and maiden full of grief, who had death in your suffering" - V-Arg1 Arg2 Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an Oblique Object (Arg3).

[N] – *they may haue grace þour wyll of the*
they may have grace though will of yours

"they may have grace though (it is) the will of yours" - Arg1 V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2).

7c(e) – 'to have (a beginning, an end, an origin)'
[N] – *vnto þat ioye whyche nevyr schal have ende*
unto that joy which never shall have end

"unto that joy which shall never have an end" - Arg1 Arg1' V Arg2

Observation: This reading requires a Subject (Arg1) and a Direct Object (Arg2). As there is a relative clause, Arg1 is duplicated as Arg1'.

[N to-INF(haven)] – *sende vs grace here to haue so good an ende*
send us grace here to have so good an ende

"send us grace here to have such a good end" - Arg1 V(to-INF(haven)) Arg2

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2) and an infinitive clause with the verb *haven* and the verbal particle 'to' (to-INF(haven)).

11(b) – 'to be under obligation (to do sth.)'
[N INF PP] – *what shulde I more mane haue do for þe*
what should I more man have do for you

"what more should I have to do for you, man?" - Arg2 Arg1 V INF Arg3

Observation: This reading requires a Subject (Arg1), a Direct Object (Arg2), an Indirect Object (Arg3) and an infinitive clause (INF).

[to-INF] – *lecchery clennesse hat mad to fle*
lust cleanness had madly to flee

"[... mercy for mankind is put away,] cleanness of lust had to flee madly" - Arg1 V to-INF

Observation: This reading requires a Subject (Arg1) and an infinitive clause with the verbal particle 'to' (to-INF).

UNCERTAIN (in the passive voice)
[ADJ] – *to save mankynd ellis al the world had be forlore*
to save mankind else all the world had be forfeit

"to save mankind all the world had to be forfeit" - Arg1 V(haven) ADJ

Observation: This reading requires a Subject (Arg1) and a past participle functioning as an Adjective (ADJ).

Q – List 8. Readings of *haven* for the 15th century according to the MED ordered into the most to the least frequent ones considering the sampled lexical occurrences.

4c(e) - 'show mercy to (sb. or sth.), take pity on' - 10 occ.
4a(a) - 'to enjoy (a right or privilege), have (the right to possess sth.); have (power etc); have (leave or license to do sth.)' - 4 occ.
4d(a) - 'to experience (a feeling, fear, pleasure, wonder, joy, etc.)' - 4 occ.
4b(a) - 'to have (sth. that denotes a state or condition)' - 3 occ.
4d(c) - 'to consider; ~ remembraunce, remember (to do sth.); ~ minde, to remember (sth.), consider' - 3 occ.
7c(a) - 'to obtain (help, peace, mercy, favor, victory, etc.)' - 3 occ.
1b(a) - 'to have (sb. or sth. somewhere)' - 2 occ.
1c - 'to possess (sth. together with the capacity to do sth. with it)' - 2 occ.
4b(c) - 'to have (cause or reason), have (reason to do sth.)' - 2 occ.
4c(c) - 'to have (a function, power, etc.) (~ sight, be able to see; ~ lif, have life, be alive)' - 2 occ.
4c(d) - 'to have (a moral or spiritual quality)' - 2 occ.
7c(e) - 'to have (a beginning, an end, an origin)' - 2 occ.
11(b) - 'to be under obligation (to do sth.)' - 2 occ.

HAPAX LEGOMENA

1a(a) - 'to possess'
1d(a) - 'to have (sb. or sth. in a state or condition)'
2(a) - 'to be provided with (a part, an organ)'
3(e) - '~ evening (felaue, per), to have an equal'
4a(g) - '~ place, of a medicine: have a proper time to be used'
4b(b) - 'suffer (pain, harm, a blow, a wound); be possessed by (devils)'
4b(a) & 4b(b) 'to have (a disease)' & 'to have (hostility)'
4c(a) - 'to have (a nature, quality, characteristic, etc.)'
4d(b) - 'to have (an idea, a thought), have (knowledge, etc.)'
6a(a) - 'to regard (sb. or sth. in a certain way), consider'
6a(c) - '~ in hate, hate (sb.)'
uncertain

R – Sum up of all readings for all covered centuries

READING/CENTURY	12	13	14	15
1a(a)	**x**	**x**	**x**	**x**
1a(e)		x	x	
1b(a)		x		x
1c	**x**	**x**	**x**	**x**
1d(a)				x
2(a)		x	x	x
2(c)			x	
2(d)	x	x	x	
3(a)	x		x	
3(b)		x	x	
3(c)	x		x	
3(d)	x	x		
3(e)		x	x	x
4a(a)			x	x
4a(e)	x			
4a(g)				x
4a(h)			x	
4b(a)	**x**	**x**	**x**	**x**
4b(a) & 4b(b)				x
4b(b)		x		x
4b(c)				x
4b(d)		x		
4b(f)	x			
4b(g)	x			
4c(a)	x			x
4c(b)		x		
4c(c)		x	x	x
4c(d)	**x**	**x**	**x**	**x**
4c(e)	**x**	**x**	**x**	**x**
4d(a)	**x**	**x**	**x**	**x**
4d(b)				x
4d(c)	x		x	x
4d(e)	x		x	
5b(a)			x	
6a(a)		x	x	x
6a(c)				x
7a(a)		x	x	
7a(c)		x		
7a(d)			x	
7b(d)	x			
7c(a)	**x**	**x**	**x**	**x**
7c(b)	x	x	x	
7c(d)			x	
7c(e)				x
9(b)		x	x	
10(b)			x	
11(b)	x		x	x
N.C. - T1		x		
N.C. - T2		x		
uncertain				x

S1 – Diatheses with their readings for the 12th century's sampled lexical occurrences.

	DIATHESES	READINGS
	[N]	1a(a); 2(b); 2(d); 3(a); 3(c); 3(d); 4b(a); 4b(f); 4c(a); 4c(d); 4c(e); 4d(a); 7b(d); 7c(a) and 7c(b)
	[N to-INF]	1c
	[N ADJ]	2(b)
	[N PP]	3(c); 4a(e); 4b(a); 4d(a); 4d(c); 7c(a)
	[V]	4b(g) *particular case
	[N for-to-INF]	4c(a)
	[N INF]	4d(a)
	[PP INF]	4d(e)
	[to-INF(haven) N]	7c(a)
	[for-to-INF(haven) N]	7c(a)
	[to-INF]	11(b)
TOTAL	**11**	**21**

S2 – Diatheses with their readings for the 13th century's sampled lexical occurrences

	DIATHESES	READINGS
	[N]	1a(a); 1b(a); 2(a); 3(d); 4b(a); 4c(b); 4c(c); 4c(d); 4d(a); 7a(a); 7a(c); 7c(a); 7c(b) and N.C. (Type 1)
	[N INF]	1a(a)
	[N PP]	1a(e); 1b(a); 2(d); 4b(a); 4b(b); 4b(d); 4c(e); 9(b) and N.C. (Type 2)
	[N to-INF]	1c
	[]	3(e)
	[ALC]	3(e)
	[N to-INF(haven)]	4b(a)
	[N INF(haven)]	4c(d) and 7c(a)
	[N CL]	6a(a)
	[CL]	7c(b)
TOTAL	**10**	**24**

S3 – Diatheses with their readings for the 14th century's sampled lexical occurrences.

	DIATHESES	READINGS
	[N]	1a(a); 2(a); 2(d); 3(a); 4a(a); 4b(a); 4c(c); 4c(d); 4d(a); 7a(a); 7c(a); 7c(b); 7c(b) and 11(b)
	[N PP CL]	1a(e)
	[N PP [CL]]	7a(d)
	[N to-INF]	1c and 11(b)
	[N for-to-INF [PP]]	1c
	[N PP for-to-INF]	2(c)
	[N PP]	3(c); 4a(h); 4c(e); 4d(a); 4d(c); 5b(a) and 6a(a)
	[ALC]	3(e)
	[N INF(haven) PP]	4c(c)
	[PP CL]	4d(e)
	[to-INF(haven) PP]	7a(a)
	[N to-INF(haven)]	7c(a)
	[N ADJ]	7c(d) and 10(b)
	[N to-INF PP]	9(b)
TOTAL	**14**	**29**

S4 – Diatheses with their readings for the 15th century's sampled lexical occurrences.

	DIATHESES	READINGS
	[N]	1a(a); 2(a); 4b(a); 4ba(a) & 4b(b), 4c(a); 4c(c); 4c(d); 4c(e); 4d(a); 4d(b); 7c(a) and 7c(e)
	[N PP]	1b(a); 1d(a); 4b(a); 4b(b); 4c(e); 4d(c); 6a(c) and 7c(a)
	[N to-INF]	1c; 4a(a); 4a(g); 4b(c)
	[]	3(e)
	[N to-INF(haven) PP]	4c(c) and 4d(c)
	[N to-INF(haven)]	4c(d) and 7c(e)
	[N CL]	4d(a)
	[N PP to-INF]	4d(c)
	[N PP ADV]	6a(a)
	[N INF PP]	11(b)
	[to-INF]	11(b)
	[ADJ]	Uncertain *passive
TOTAL	**12**	**25**