1. Introduction

When discussing how to prepare remedies, medieval remedybooks featured plant ingredients, but also those of other substances. It is these substances which this study examines. Despite the claim that "mineral and chemical elements are unusual" in medical recipes (Hunt 1990: 19), our research reveals that a good number of chemical terms are, in fact, present in Middle English medical manuscripts. Such findings accord with Norri's comment that the "study of untapped manuscript material would bring to light a vast number of words and phrases unrecorded in any of the historical dictionaries of English" (2016: 9). Indeed, some of the items in our corpus could not be identified. Thus, our aim is to study the chemical ingredients in late medieval English medical manuscripts to explore a field which has not attracted enough scholarly attention.

In fact, we follow Norri's work (1991) being of the few ones that deal with this topic in Middle English medical manuscripts. Thus, the delimitation of the set of *chemical* items follows Norri's classification closely (1991: 216-217). According to him, the names of chemical substances comprise four groups: the first one is made up of units that are "man-made medicinal ingredients", such as *ale*, *sope* and *tarr*. The second category consists of minerals and their preparations, like *gold*, *leed*, *alum*, *arsnyk*. Norri warns the reader about the fact that this category overlaps the first one partly, since "compounds of metal are either man-made (e.g. *white lede*, *red led*) or occur freely in nature (e.g. *arpiment*, *resalger*)" (1991: 217). The third group includes extracts from plants, such as gums, resins, juices and dyes. Finally, the last category has to do with terms containing "medicinal earth" (Norri 1991: 2165), such as *bol(e armonyak* and *terra sigillata*. To sum up, here *chemical* is used as an umbrella term to refer to substances such as metals and their compounds, extracts from plants, medical earth and man-made medical ingredients.

After the introduction, the methodology section concentrates on the explanation of the compilation of the corpus and the problems encountered when delimiting and examining the chemical field. This is followed by the analysis of the data. Both a quantitative and a qualitative examination is provided. We discuss the linguistic sources of the nouns, as well as the structure and the constituents present in nominal compounds according to the usual taxonomies based on Bauer (1983 and 2017), Kastovsky (1992) and Marchand (1969), but specialised classifications on the topic are also used (Norri 1991). This section is followed by a comparison between the findings in the field of chemical nominal combinations and in botany based on a previous study we undertook on plant compounds with data extracted from a similar corpus of medical texts. Finally, the conclusions from the analysis are presented.

2. Methodology

To carry out the study of medieval chemical vocabulary, a corpus of late Middle English medical texts has been specially compiled from different libraries, chiefly Glasgow University Library (henceforth GUL), both from the Hunter and the Ferguson Collections, the British Library and the Wellcome Library. We have tried to cover several genres included within the classification by Pahta and Taavitsainen (2004: 15). Thus, our corpus, which includes approximately 215,000 words, contains fourteenth-century and fifteenth-century medical texts categorised as:

1) Specialised treatises: a humoral tract in British Library, Sloane 121 and the Middle English fifteenth-century translation of the *Compendium Medicinae* by Gilbertus Anglicus in Wellcome 537.

2) *Materia medica* and remedybooks which comprise both a) Herbaries and other related works: GUL Ferguson 147 *Antidotarium Nicholai*; GUL Hunter 185 *Flora medica*; GUL Hunter 307 *Pharmacopoeia* and a herbary in Hunter 497[2] and b) Recipe collections: GUL Ferguson 147 and GUL Hunter 328 *Alphabetical List of Medicines* and *Alphabetical List of Remedies*.

Once the texts were fully transcribed, the chemical ingredients were identified in each of the manuscripts. Following Norri's classification (1991: 216–217), metals and natural elements like gums, resins and earthly substances, but also man-processed items are included, for instance oils, juices, flour, ale, wine and tar. We have incorporated man-processed products that could be considered simple remedies, but have disregarded others that were made of several ingredients, such as electuaries or ointments, even if the main ingredient was a mineral or a substance, such as a resin. The reason behind this decision is the fact that both electuaries and ointments can be considered compound medicines, since, according to the *Oxford English Dictionary* (henceforth *OED*), they are defined as:

> Electuary: "A medicinal conserve or paste, consisting of a powder or other ingredient mixed with honey, preserve, or syrup of some kind".
>
> Ointment: "An unctuous preparation, of a soft consistence like that of butter, often mixed with some medicament".

The identification of the data has been carried out by using different lexicographic works available to us: chiefly the *OED* and the *Middle English Dictionary* (hereafter *MED*), but also Hunt (1989), Getz (1991) and Norri (2016). Before attempting the classification of the data, it is worth clarifying our use of foreign names.

Norri (2016: 5) acknowledges the difficulty of deciding whether a highly technical word adopted from a foreign language can be said to be common or unusual. Likewise, scholars working on code-switching cannot determine with certainty when a word is a borrowing or an instance of code-switching. Subsequently, out of the three hundred and forty-five attested chemical items, it is difficult to ascertain accurately how many are foreign words and how many are English words. Loanwords are usually considered to be somehow integrated into the linguistic system and adapted to its spelling, morphology, etc., whereas foreign elements are not. In the case of chemical lexicon, it is not always easy to determine whether they belong to one type or the other. This is primarily due to the fact the phenomenon of code-switching is very common in medical texts.[3] The scribe often integrates Latin within the English text in a very subtle way. However, for the analysis we classified several lexical units in a specific category that were clearly in Latin: *olium popilium*, *olium ederatum*, *olium liliacum*, *olium mandragoratum*, *olium mirtum*, *olium pulegum*, *olium sambuonum*, *olium sanbuci*, *olium savinum* and *olium vrolanum*, among others. Likewise, French combinations, like *frankencens*, *sangdragon*, *sandeuer*, *verdegres* and *viueh chauh*, also fall into the category of foreign elements.

2 The *Antidotarium Nicholai* in Ferguson 147 and the manuscript Hunter 185 were edited and studied by Carrillo-Linares (1997) and Alonso-Almeida (2000, 2014) respectively. In addition, *De Viribus Herbarum* contained in Hunter 497 was edited by Miranda-García and Calle-Martín (2012).

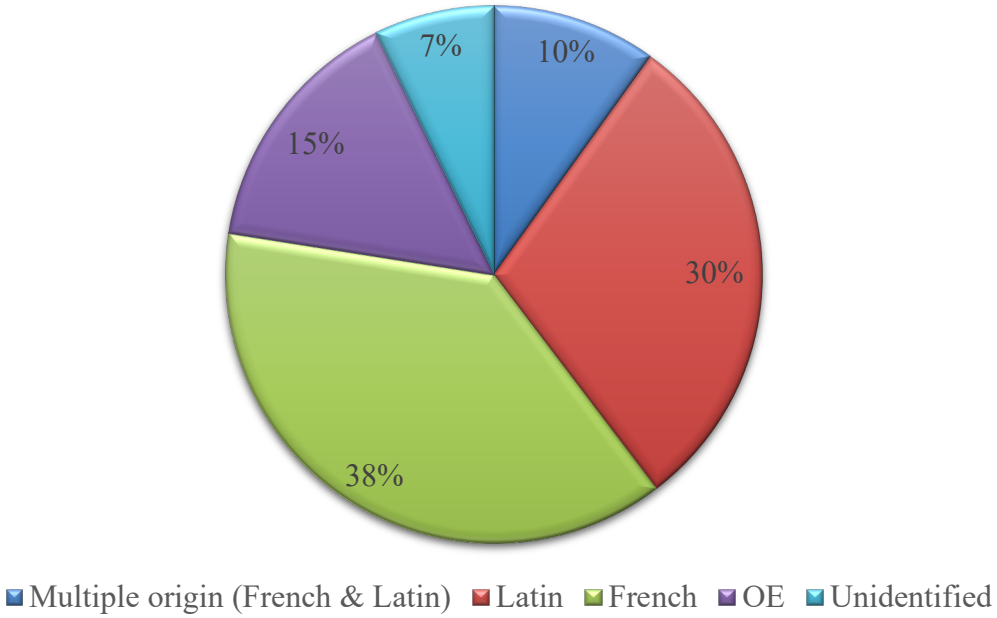
3 The influence of Latin on English medical texts has been specifically evaluated by Pahta (2004), while the phenomenon of code-switching in historical texts in general has been dealt with in Skaffari (2016), Skaffari and Mäkilähde (2014) and Schendler and Wright (2011), among others.

3. Data analysis

All the lexical units that have been extracted from the corpus belong to the category of nouns. A first broad classification regarding chemical terminology distinguishes between *simplex* terms and noun combinations.

3.1. Simplex terms

One hundred and eleven *simplex* terms have been identified in our corpus. Theirs origins have been analysed by taking the information provided by the *OED* in the first place. When the noun has not been found in this source, the etymological information offered by the *MED* has been considered.[4] Moreover, when the word appears in both sources, the information has been collated giving always priority to the *OED*. Sometimes glossaries have also been used, specially the one carried out by Getz (1991), as well as Norri's *Dictionary of Medical Vocabulary* (2016). The distribution of *simplex* nouns can be seen in Graph 1.



Graph 1. Etymological origin of simplex nouns[5]

As Graph 1 reflects, *simplex* terms coming from Old French stand for the largest group in our corpus. Forty-two examples have been recorded, a number which represents approximately 38% of the total. Among all the nouns identified within this group, it is interesting to draw special attention to several items, since the information provided by the *OED* and *MED* regarding their etymological origin does not always coincide:

4 In fact, *The Dictionary of Medical Vocabulary* by Norri (2016) has also been systematically consulted. Since most items were wanting or provided no new information on the specific items, this dictionary was finally disregarded to be considered at the same level as the *OED* and the *MED*.

5 The real percentages in descending order are French (37.83%); Latin (29.73%): OE (15.31%); Multiple origin (9.91%); and unidentified (7.20%). All figures have been rounded off to the nearest whole number.

1) *Musk*: the *OED* claims that it is a borrowing from French. This source defines the term as "a reddish-brown substance with a strong, persistent odour secreted by a gland of the male musk deer, esp. *Moschus moschiferus*, and highly prized in perfumery. Also: any of various substances secreted by other mammals, esp. for scent marking". This definition is perfectly in line with Getz's glossary, where the noun appears as "secretions of the preputial follicles of the musk deer" (1991: 343). Substances of an animal origin have not been included in this study. However, when these definitions and sources are complemented with the information taken from the *MED*, a second meaning that is not present in the other sources is found: "some kind of sap or juice obtained from various plants; musk of trees, musk of the water lemke". The consultation of both lexicographic works helps to complete the actual semantic area covered by the term in Middle English. When referring to the alchemical field, Grund states that both dictionaries provide researchers with "complementary pictures rather than a unified picture of alchemical words and meanings, and they seem best consulted in tandem rather than in isolation" (2013: 594). This statement perfectly applies to medical lexicon.

2) The etymological information given by the two sources that have been used to identify and define the *simplex* terms does not always coincide. As far as the noun *galle* is concerned, the *MED* claims it comes from Latin *galla*, whereas the *OED* states it comes from French *galle*. It is highly likely that the French term came from Latin. Nevertheless, this mismatch can be illustrated with another example. In the etymological information provided by the *OED* for *gumme*, it reads: "< Old French *gomme* = Provençal *goma*, Spanish *goma*, Portuguese *goma*, Italian *gomma* < popular Latin *gumma* = classical Latin *gummi*, *cummi*, < Greek *κόμμι*". While the *OED* clearly assigns the item a French origin, the *MED* considers that it goes back to both languages, Old French and Latin.

3) There is one example in which the product is referred to by two different *simplex* nouns that enter English from two different languages: v*inegre* and *eysel*. The former was adopted from Latin whereas the latter was introduced through Old French *aisil*, *aissil*, which corresponds to late Latin **acētillum*, diminutive of *acētum* "vinegar", according to the *OED*.

As for the nouns coming from Latin, there are thirty-three, representing approximately 30% of the *simplex* terms analysed in this subcorpus. The great majority of these Latin words come from Greek, such as *sarcocolla* (*OED* "late Latin, < Greek σαρκόκολλα"). It is also possible to identify here some terms, such as *mirre* where the *MED* and the *OED* do not coincide from an etymological point of view. While the former points out that the nouns come from "OE myrra, myrre (from L) & L myrrha, murra & OF mirre", the second source directly says that it is "a borrowing from Latin".

As explained at the beginning of the study of *simplex* terms, the information provided by the *OED* has been given priority, even when it does not agree with the *MED* or our own data. Thus, the *OED* considers *sarcocoll* rare today and documented for the first time in *Lanfranc's Science of Cirurgie*, *circa* 1400 "Take oile of rosis..mirre, sarcocol". Nonetheless, our corpus records the item as "Sarcocolla is a gumme. h . d . in . j ." in GUL, Hunter 307, a fourteenth-century manuscript. This means the term is attested in our corpus before the date provided by the *OED*. Another example of antedating is found in Wellcome 537 (f. 88r) where the reader is advised to mix the following ingredients to get a collyrium: "Anoþir colliry: take roses, cathmie, gumme arabik, dragagantum, amide, ana, dr. iiii." (Getz 1991: 46). The term *cadmia* is defined by the *OED* as "The ancient name of calamine' (Ure Dict. Arts I. 569); also applied to a sublimate consisting of oxide of zinc (tutty), and to an ore of cobalt" and recorded for the first time in 1657. Nonetheless, Wellcome 537 has been dated by Getz "before about 1460" (1991: lxii).

With regard to nouns with a Germanic origin, there are seventeen items, which represents around 15% of the words studied in this subcorpus of English nouns. All these terms appear in the *OED* as "words inherited from Germanic". However, the *MED* states explicitly that all of them come from Old English. In this section there is one element, *talowe*, which deserves special attention. It is controversial not only regarding its origin, but also in terms of its meaning and subsequent inclusion in this study.

Talowe has been identified in Wellcome 537 and therefore appears in Getz's glossary (1991: 357). It is defined here as "rendered fat", which perfectly coincides not only with the definition provided by the *MED* "The fatty tissue of an animal; fatty tissue from around the kidneys or other internal organs of an animal", but also with two meanings offered by the *OED*: "1) The fat or adipose tissue of an animal; 2) A substance consisting of a somewhat hard animal fat (esp. that obtained from the parts about the kidneys of ruminating animals, now chiefly the sheep and ox), separated by melting and clarifying from the membranes, etc., naturally mixed with it; used for making candles and soap, dressing leather, and other purposes". Additionally, the *OED* offers a third meaning: "Applied to various kinds of grease or greasy substances, e.g. those obtained from plants". Due to the fact that it can also be a substance which derives from plants, it has been considered for this study.

As for its origin, on the one hand, the *OED* claims it is recorded in "Middle English *talȝ*, *talgh*, known first in 14th cent.; corresponds to Middle Low German *talg*, *talch*, Low German *talg*, in early modern Dutch *talg*, *talch* (16th cent.), Dutch *talk* feminine and German *talg*, in 1572 *talck* masculine; Modern Icelandic (14th cent.) *tólg*, *tólk*, Middle Danish (13th cent.) *talgh*, *talwh*". On the other hand, the *MED* attested a protoform in Old English "**tealg*(-, **talg*(-: cp. MLG *tallich*, *talch*, MDu. *talch*, older Dan. *talgh*, & G *Talg*; & cp. ML *talo tallow*". In this case, the information from the *MED* has been a key factor to classify this term under the label of Old English.

Regarding the *simplex* terms that have been classified in the group "multiple origin: French and Latin", eleven examples have been identified, which represents the 10% of *simplex* nouns. Three of them have only been found in the *MED*: *dragagantu*(*um*), *tuchia* and *sarapinum*. In the case of *sarapinum,* its etymology according to this dictionary is ML *serāpīnum* and OF *serapin*. Then, there is only one instance of an item which in the *OED* appears as "either (i) a borrowing from French. Or (ii) a borrowing from Latin", that is the noun *maces*. The origin of the remaining nouns (*acacie*, a*sernici*, m*astik*, *opoponac*, *prassin*, *resyna* and *sandalis*) is defined by the *OED* as "of multiple origins. Partly a borrowing from French. Partly a borrowing from Latin".

Finally, there is a small amount of terms which have not been successfully identified. The great majority of them appear in Wellcome 537 and have been taken into consideration since the definition offered by Getz (1991: 339 and 344) suits the guidelines established for this study. Therefore, nouns such as *onyfacie* ("juice of green white grapes") or *macematicon* ("cabbage juice") have been included. However, this glossary contains no information regarding their origin, and they have not been found in the *OED* nor in the *MED*. There are eight items under the label *unidentified*. This is probably due to the fact that, as Grund (2013: 594, 595) pointed out for the alchemical field, "numerous lexemes and meanings still remain unrecorded", although "the *MED* and the *OED* contain a wealth of information on alchemical vocabulary in English from the fourteenth and fifteenth centuries".

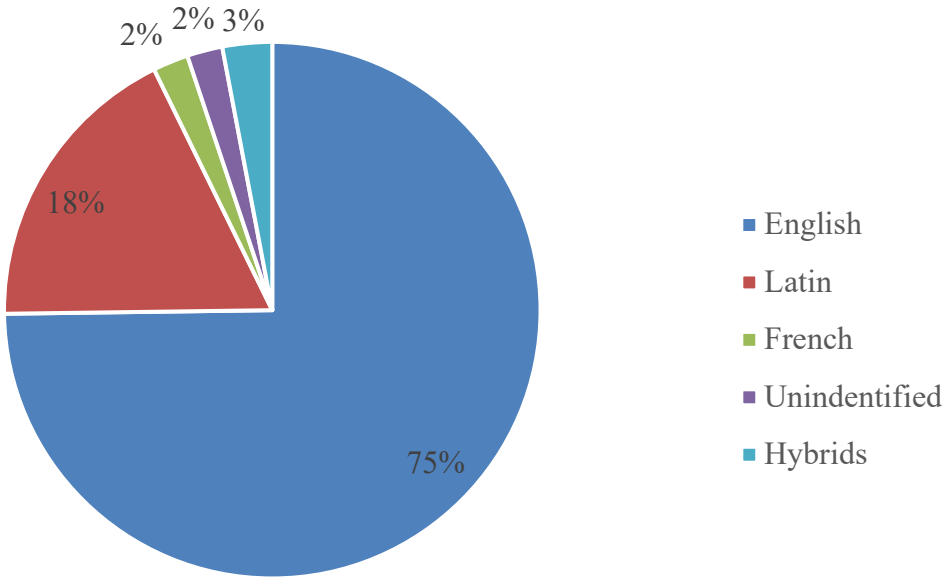
3.2. Noun Combinations

First of all, a distinction between noun combinations in English and other languages, such as French and Latin, needs to be made. Even though foreign combinations contribute to picture the map of the scientific terminology of the time, they have not been coined in English, which is the focus of our analysis. As can be seen in Graph 2,

the Latin formations are 18% of the total. Most Latin compounds are made of a noun followed by an adjective, as in *lapis lazuli*. We will not get into further detail regarding the Latin compounds, since, even though their presence depicts clearly the mixture of languages found in scientific English in the period, they do not give real data on the linguistic situation of the most used structural patterns in the English language. Unlike Latin, French records only five items: *frankencens*, *sang dragon*, *sandeuer*, *verd gres*, *viueh chauh*. Similarly, we find some hybrids, where part is in English and part in French or Latin, such as in *oyle de bay, oyle of Iuniperi*. None of these will be considered for the analysis.

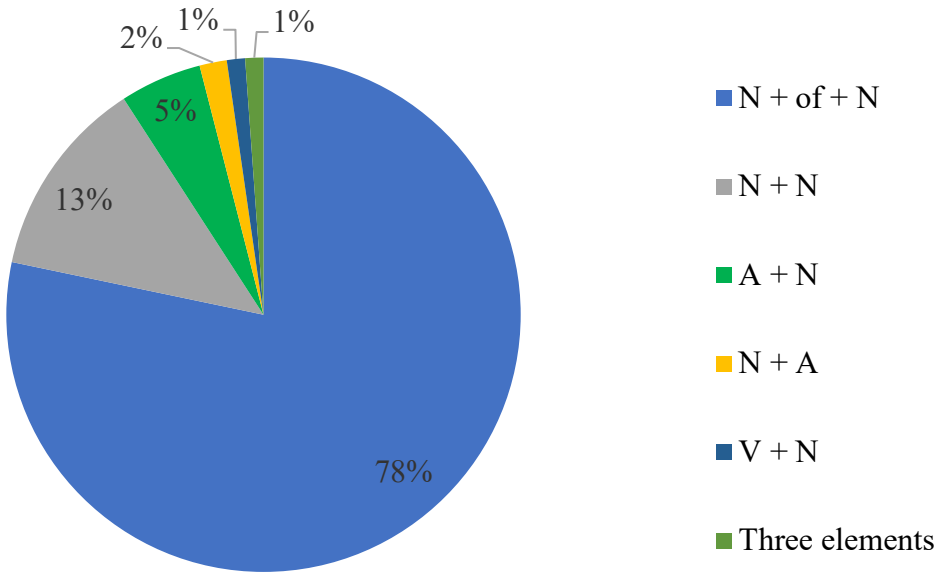
Finally, there is a small group of combinations that could not be identified. For instance: *cakstorax calamyte*, *calen viriden*, *ellus nigrus*, *lyridlyme* and *sarcoll bedilly*. Regarding the first one, a similar combination, *storax calamyte*, was identified and classified as English compound, but the first element in *caskstorax* could not be found. Subsequently, the grammatical category is unknown: whether it can be a verb, a noun or an adjective followed by the noun *storax*. A similar case is *calen viriden*; the last part of the compound appears in the *OED* as *viridin* from French meaning "chorophyl", but the first element could not be identified. Regarding *ellus nigrus*, the origin seems to be Latin, since it appears in a sequence with other *chemica* ingredients in Latin: "lapis laȝuli, lapis armonicus, fumus terre, lapis magnetis, lapis lincis, ellus nigrus, lapidus calcis" (GUL, Hunter 185). The actual spelling of the first element is *ell*^{9} with a superscript abbreviation which has been interpreted as *-us*. We wonder whether this form could be an abbreviation for *elementus* or even *elleborus*, "any of various medicinal plants valued chiefly for their strong purgative properties, mostly belonging to the genera Veratrum (family Melanthiaceae) and Helleborus (family Ranunculaceae)" (*OED*). Whichever reading was meant by the scribe, no identification of either *ellus nigrus* nor *elementus nigrus* has been possible. This means it cannot even be assured that *ellus nigrus* is a mineral and not a plant, as could be the case of *elleborus nigrus*. The same applies to *lyridlyme* and *sarcoll bedilly*.

Thus, the distribution of English, foreign and unidentified noun combinations is presented in Graph 2.



Graph 2. Proportion of English, foreign and unidentified combinations

The main focus of interest is on the 75% of the corpus which corresponds to English noun formations in Graph 2. Here the taxonomy and description of nominal structures will be based on word-class affiliation of the *determinatum* or head of the combination (Kastovsky 1992: 365). Thus, in this study they are divided into groups according to the parts of speech they are composed of. Numerically speaking, the distribution of the different types found can be seen in Graph 3.



Graph 3. Percentage of English noun combinations[6]

When concentrating on the category of Noun + *of* + Noun, a word of warning should be said as regards the inclusion of this group in the analysis. Even if experts admit that "what counts as a compound may be fuzzy" (Bauer 2017: 1), these Noun + *of* + Noun structures can hardly be considered compounds but free combinations. Nevertheless, previous research on nominal formations on minerals (Norri 1991: 231) and plants (Sauer 1995: 314) have included them in their studies. Here, we mainly find juices, oils and syrups. *Juice* with various alternant spellings (*ois*, *yois*, *ius*, *jus*) accounts for the biggest part of the items (eighty-six). The second largest group within this category is different kinds of oils, which comprise twenty-eight items, followed by syrups which include fifteen combinations.

Apart from oils, juices and syrups, the rest of the combinations refer to other substances, such as gums like *gum(me) of yuy*; metals, such as *lymail of bras*, *lymail of gold*, *lymail of siluer* and *lytarge of leed*, which according to Norri "seems at first a tautologous formation, as *litarge* is used for lead monoxide usually without a modifier" (1991: 231). Finally, flours such as *mele of cokkul*. The latter is defined by the *OED* as a name "applied to Lychnis (or Agrostemma) Githago, a caryophyllaceous plant, with handsome reddish-purple flowers

6 The real percentages in descending order are: N + of + N (78.28%); N + N (12.57%): Adj + N (5.14%); N + Adj (1.71%); V + N (1.14%); and three element compounds (1.14%). All figures have been rounded off to the nearest whole number.

succeeded by capsules of numerous black seeds, which grows in cornfields, especially among wheat". Wines, like *wyne of pomganard* also appears in the recipes as well as *tarre of ynde*.

The second largest category is made of a Noun + Noun. Here we find lexical units such as *ademant stoon*; *barly mele*; *bene mele*; *bere mele*; *boles galle*; *cokyl mele*; *golde foyle*; *gum(me) edere*; *houstuus mele*; *loodestone*; *margery perles*; *mel roses*; *mel roset*; *mete oyl(e)*; *oote mele*; *oyle roset*; *rye mele*; *salt alkaly*; *salt geme*; *siluer foyle*; *sugre roset*; *whete mele*. As can be seen, most of them are flours from different cereals, although salt-like substances in which the word salt is postmodified are also found, as in *salt alkaly* and *salt geme*.

The third group that ranks high is Adjective + Noun. Here combinations, such as *arabike gummi*; *comynne oyle*; *comune salt*; *lasse mirre*; *quicksilver*; *rede lede*; *rede wyne*; *white wyn*; *whyte lede*, are recorded. The connection between some of these elements seems to be firmly established. Both *arabike gummi* and *gumme arabik* with inversion in the order of elements are frequently found. Likewise, *red wine* and *white wine* are common ingredients in many recipes as well. *Quicksilver* used to be employed to heal common ailments. Here *quick* refers to "living", as stated by the *OED*. Along with other chemical substances, it is one of the ingredients of the following unguent for *Serpiginosis and Elefanciosis*:

> Take aloes lytarge arsenik quicsiluer argule mastic olibanum comyn soufre vif pyche olde gres þilke þat beþe to poudrynge grynde hem, and tempre wyth vynegre and wyth þe Iois of fumterre afturwarde oyngte the pacyent by þe fuyre or ellis by þe sunne (GUL, Ferguson 147, f. 53v)

In the following category, three combinations of a Noun + Adjective are registered: *gumme arabik*, which is a resin from certain species of acacia; *bole armonyak* and *sal(t) armoniac*. The last two contain the same adjective. The former "is a corruption of *Armeniac* '*Bole Armoniac* or the *Armenian Bole* is a soft friable fatty earth, usually of a pale red colour" (*OED*), while the latter refers to "Salt of Ammon, a hard white opaque crystalline salt, supposed to have been originally prepared from the dung of camels near the temple of Jupiter Ammon, as it still is in Egypt; chemically Ammonium Chloride NH_4Cl, formerly called muriate of ammonia; used in tinning iron, in pharmacy, and for the manufacture of Ammonium Alum for the dyer" (*OED*).

The final categories contain two items each: Verb + Noun is documented in *brimstone,* where the Old English verb *brinnan* "to burn" plus the noun *stone* are used to mean "sulphur". The other element in this category is *whetston*. Regarding the formations with three elements, two different combination patterns are found, one with three nouns, as in *oyle roset cense* and another one, made up of a noun followed by an of-phrase where the head is preceded by another noun in genitive: *syrip of smyþþis water*. The *OED* defines *smith's water* as "water in which a smith has cooled hot iron". The three-element combination is found in Wellcome 537, as well as in Hunter 95 where a recipe can be found for a *Cankyr of the eiȝe*, where eye drops are prescribed from "ceruse dragage . folium roses of eueriche j lyche myche" and further down the reader is assured "smyþis water is profitabil . for þis greuaunce" (Hunter 95, f. 41r).

4. Comparison with plant combinations

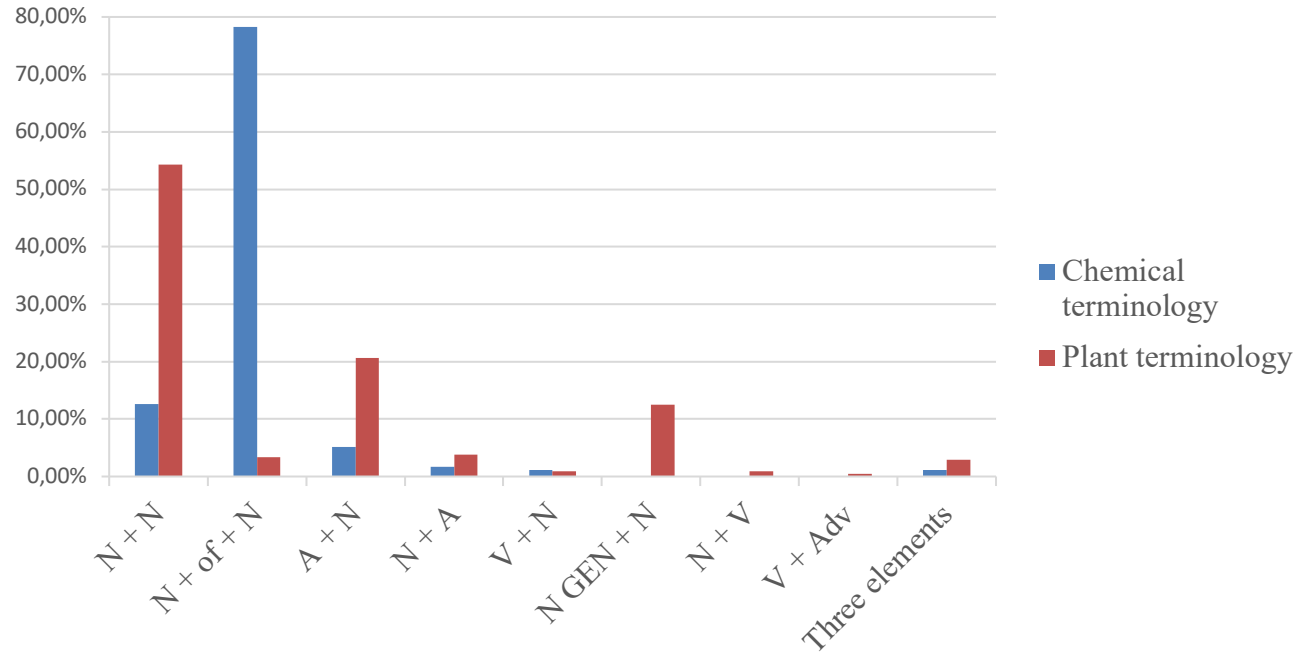
No quantitative analyses of chemical terminology have been carried out in Middle English thus far. However, there are some statistical studies on Old English (Krische 2010) and Middle English (De la Cruz-Cabanillas and Diego-Rodríguez 2018) plant names in terms of noun combinations. Thus, in a previous study on the classes of noun formations found in Middle English botanical terminology, the following categories were identified:

1. **Noun + Noun**: *hors mynte*; *ribwort*

2. **Adjective + Noun**: *whyte peper*; *wildeneep*
3. **Noun in GENITIVE + Noun**: *woselystonge*; *faytores herbe*
4. **Noun + *of* + Noun**: bene of fraunce; peletre of spaigne
5. **Noun + Adjective**: *hollyhock*; *rose marine*
6. **Verb** + **Adverb**: *carewey*
7. **Verb + Noun**: *standelegursse*
8. **Noun + Verb**: *wodebynde*
9. **Three elements**: *herbe seint Iohn*; *oure lady þistil*

Since the real number of items in each corpus was different, the percentages are shown for each series: first, the chemical noun combinations and second, the plant noun combinations. By having a look at the taxonomy, it can be realised that three structural patterns are missing in our study on chemical ingredients: Noun in GENITIVE + Noun, Verb + Adverb and Noun + Verb. Even if Bauer states (2017: 24) that "there is no hard line between N_{GEN} + N and N + N constructions", the results in both corpora are remarkably different. The outcome in plant names is in line with previous studies on the topic, but the chemical field seems to show preference for the Noun + *of* + Noun combinations. Sauer (1995: 314), when referring to plant names states that "this is a French pattern which did not exist in OE, but was taken as a loan pattern in the course of ME", In fact, Krische (2010) in her detailed study of Old English plant terminology does not include this combination. The fact that it is the most productive structure in our corpus with 78.28 % of the total may be the result of the high presence of Latin and French terminology in this field. The combination was calqued in the Middle English period through translations of classical texts or the copy of foreign patterns as a result of the vernacularisation process that took place during the fourteenth and fifteenth centuries. Pahta and Taavitsainen (2004: 11) claim that academic and surgical treatises, being new in the vernacular, "were translated or adapted from Latin sources". Furthermore, they add that "most Middle English scientific texts are translated or, in one way or another, derived from Latin or French treatises" (2004: 13).

The visual distribution of each field can be seen in comparison with the other one in Graph 4.



Graph 4. Comparison of chemical and plant noun combinations

Other findings related to the comparison of plant and chemical terminology show relevant mismatches as well. For instance, the Adjective + Noun construction ranks very low in chemical items (5.14%), whereas it is the second group in the plant names (20.67%). The other categories also present divergences, since Noun + Noun is the first combination in plant names with 54.33%, while in the chemical corpus it is only 12.57 %. Other constructions show slighter differences. Thus, the Verb +Noun formation represents 1.14% in the chemical corpus versus 0.96 % of the total in plant names. Likewise, the three-element combinations are 1.14% in the chemical corpus, while they are 2.88 % in the plant corpus. Nevertheless, the presence of these two formations in each of the subcorpora is not significant.

5. Conclusions

The study of chemical substances in Middle English is a vast unexplored territory where very few pioneers have dared to venture. In the previous pages, a detailed analysis of chemical ingredients present in late Middle English medical texts has been carried out. In order to do so, a purposely built corpus of approximate 215,000 words was compiled and transcribed when needed. The chemical substances were identified and classified into *simplex* and noun combinations. This is the first systematic examination of the chemical material, both from a quantitative and qualitative perspective. Before this article very few studies dealt with the topic and no complete analysis of the lexis from the quantitative nor from the qualitative perspective had been undertaken.

The classification of nouns is based on whether the denomination is made up of only one word (*simplex*) or if it consists of a combination of two or three words. The latter group, formations of two words mainly, has been considered for the analysis, even if on some occasions there may be doubts whether they can stand for a free combination. The taxonomy used for the classification is based on word-class affiliation of the *determinatum* or head of the combination.

The findings show that the most productive structure is the Noun + *of* + Noun formation with 78% of all English combinations. Explanations for the predominance of this pattern can be found in the original Latin and especially French sources. The second most frequent combination consisted of Noun + Noun with 13% of the total. The other combinations (Adjective + Noun, Noun + Adjective and three-element word formations) account for only 9% of the total items.

The comparison with the plant name terminology reveals that even though there are coincidences in the type of structures found in both corpora, some of them are only attested in plant names. Here, the most salient group is that of Noun + Noun compounds, although this category ranks much lower in the chemical field, where the main recurrent pattern is the one mentioned above (Noun + *of* + Noun).

Thus, the use of untapped material has served to show the real picture of the chemical wordstock in Middle English medical texts. Yet there is ample space for future research, especially in the field of alchemical material, which remains largely unexplored and contains a wide array of chemical substances. This undertake will require a great deal of courage, time and energy, since very little has been done so far in this area.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

About the authors

Isabel de la Cruz-Cabanillas is a Professor of English at University of Alcalá, where she teaches several courses on English Language and Linguistics at undergraduate and postgraduate level. Her research interests are chiefly in the field of lexicology, semantics and manuscript studies. She has contributed to international journals, as well as to volumes edited by international publishing houses. Likewise, she has carried out research at the Universities of Glasgow, Sheffield, Berlin, Limerick and Oxford. Dr de la Cruz-Cabanillas serves or has served on the Editorial Boards of several journals: *SELIM*, *RESLA, Journal of English Studies, Atlantis, International Journal of English Studies, ELIA, Variants* and others.

Irene Diego-Rodríguez is a lecturer at Nebrija University. Her research interests are chiefly in the field of historical linguistics, palaeography and manuscript studies. Dr Diego-Rodríguez has contributed to international conferences, journals, and volumes edited by international publishing houses. She has also carried out research at the University of Glasgow and several British libraries.

References

Alonso-Almeida, F. (2000). *Edition and Study of a Late Medieval English Medical Receptarium: GUL MS Hunter 185 (T. 8.17)*. Unpublished Doctoral Dissertation. University of Las Palmas de Gran Canaria.

Alonso-Almeida, F. (2014). *A Middle English Medical Remedy Book*. Heidelberg: Winter.

Bauer, L. (1983). *Word Formation*. Cambridge: Cambridge University Press.

Bauer, L. (2017). *Compounds and Compounding*. Cambridge: Cambridge University Press.

Carrillo-Linares, M.J. (1997). Edición de una Versión en Inglés Medio del *Antidotarium Nicholai*. Unpublished Doctoral Dissertation: University of Sevilla.

De la Cruz-Cabanillas, I. & Diego-Rodríguez, I. (2018). Therapeutic Plant Names: Neologising, Borrowing and Compounding in a Late Middle English Medical Corpus. Presented at the *20th International Conference of English Historical Linguistics (ICEHL)*. University of Edinburgh, Scotland.

Getz, F.M. (1991). *Healing and Society in Medieval England: A Middle English Translation of the Pharmaceutical Writings of Gilbertus Anglicus*. Madison: University of Winsconsin Press.

Grund, P.J. (2013). The 'Forgotten' Language of Middle English Alchemy: Exploring Alchemical Lexis in the *MED* and the *OED*. *Review of English Studies,* 65(271), 575–95.

Hunt, T. (1989). *Plant Names of Medieval England*. Cambridge: D. S. Brewer.

Hunt, T. (1990). *Popular Medicine in Thirteenth Century England*. Cambridge: D. S. Brewer.

Kastovsky, D. (1992). Semantics and Vocabulary. In R. Hogg (Ed.), *The Cambridge History of the English Language. Vol. 1 The Beginnings to 1066* (pp. 290–408). Cambridge: Cambridge University Press.

Krische, U. (2010). *The Old English Complex Plant Names: A Linguistic Survey and a Catalogue*. Frankfurt am Main: Peter Lang.

Marchand, H. (1969). *The Categories and Types of Present-Day English Word-Formation. A Synchronic and Diachronic Approach*, 2nd ed. Munich: C. H. Beck.

Middle English Dictionary. Online edition: <https://quod.lib.umich.edu/m/med/> [20/12/2020].

Miranda-García, A. and Calle-Martín, J. (2012). The Middle English Version of *De Viribus Herbarum* (GUL MS Hunter 497, ff. 1r-92r). Edition and Philological Study. Frankfurt am Main: Peter Lang.

Norri, J. (1991). Notes on the Origin and Meaning of Chemical Terms in Middle English Manuscripts. *Neophilologische Mitteilungen*, 92, 215–36.

Norri, J. (2016). *Dictionary of Medical Vocabulary in English, 1375–1550*, Vols. 1, 2. London: Routledge.

Oxford English Dictionary. Online edition: <www.oed.com> [20/12/2020].

Pahta, P. (2004). Code-Switching in Medieval Medical Writing. In I. Taavitsainen & P. Pahta (Eds.), *Medical and Scientific Writing in Late Medieval English* (pp. 73–99). Cambridge: Cambridge University Press.

Pahta, P. & Taavitsainen, I. (2004). Vernacularisation of Scientific and Medical Writings in its Sociohistorical Context. In I. Taavitsainen & P. Pahta (Eds.), *Medical and Scientific Writing in Late Medieval English* (pp. 1–22). Cambridge: Cambridge University Press.

Sauer, H. (1995). On the Analysis of Old and Middle English Plant Names. In B. Santano-Moreno (Ed.), *Proceedings of the 10th International Conference of the Spanish Society of Medieval Language and Literature* (pp. 299–325). Cáceres: Universidad de Extremadura

Sauer, H. (2011). Patterns of Loan-Influence on the Medieval English Plant Names, with Special Reference to the Influence of Greek. In J. Fisiak, & M. Bator (Eds.), *Foreign Influence on Medieval English* (pp. 55–76). Frankfurt am Main: Peter Lang.

Schendler, H. & Wright, L. (2011). Code-Switching in Early English: Historical Background and Methodological and Theoretical Issues. In H. Schendler & L. Wright (Eds.), *Code-Switching in Early English* (pp. 15-46). Berlin: De Gruyter.

Skaffari, J. (2016). Code-Switching and Vernacular Support: An Early Middle English Case Study. *Multilingua*, 35(2), 203-226.

Sakaffari, J. & Mäkilähde. A. (2014). Code-Switching in Historical Materials: Research at the Limits of Contact Linguistics. In R. Nicolaï (Ed.), *Questioning Language Contact: Limits of Contact* (pp. 252-279). Leiden: Brill.