

Pécsi Tudományegyetem
University of Pécs

Bölcsészettudományi Kar
Faculty of Humanities

Historical Code-mixing and Hybrid Place-names in England

PhD Dissertation

Written by:

Tamás Fekete
Doctoral School of Linguistics

Supervised by:

Dr. habil. Irén Hegedűs
associate professor
Institute of English Studies
Department of English Linguistics

Pécs, 2017

Abstract

The purpose of the dissertation is to present a corpus-based analysis of hybrid place-names occurring in England within the framework of historical sociolinguistics and historical code-switching. The various types of place-name formations found in England will be presented and it will be argued that certain formations can be construed as instances of historical code-mixing and that the sociolinguistic and stratal relationship of languages involved in the creation of hybrid place-names is the main determinant of the outcome of toponymic influence. The nature of the creation of toponyms will also be surveyed, and it will be argued that they are in fact darkened compounds whose semantic content is the main determiner of the changes that will affect them.

Keywords: *place-names, onomastics, historical linguistics, corpus linguistics, Old English*

Acknowledgements

I would like to express my gratitude to...

...my supervisor, Irén Hegedűs for her continued guidance and help since I wrote my BA thesis in 2010.

...all my professors at university, especially József Andor, for their guidance throughout this ten-year-long journey since my BA studies

...my parents, Tibor and Edit for their life-long support and encouragement, for standing by me and for giving me the liberty to study what I have always been most interested in

...my friend and colleague Ádám Porkoláb for his invaluable help with the compilation of the place-name corpus and for the helpful discussions

...my friends for their encouragement and discussions

...the organizers and participants of the 1st Historical Code-switching: The Next Step conference in Tampere in June 2014, where the first version of a paper which later grew out into this dissertation was presented, I am grateful for their questions, comments, and suggestions.

Finally, I would like to thank the anonymous reviewers of the conference proceedings, scholarly journals and all other scholarly papers for their comments and suggestions concerning the works I submitted to them. Many of those suggestions have been incorporated into this dissertation. Any remaining mistakes are of course mine.

As an honorary mention, I would also like to thank Tim Berners-Lee for inventing the World Wide Web.

Historical Code-mixing and Hybrid Place-names in England

Table of Contents

I. Introduction

1.1. Introduction.....	10
1.2. Background and purpose of the dissertation.....	11
1.3. Research questions.....	16
1.4. Structure of the dissertation.....	19
1.5. Limitations of the research.....	20

II. Code-mixing and other Theoretical Considerations

2.1. The notion of code-mixing and its relevance for hybrid place-names.....	22
2.2. The theoretical framework of code-switching and code-mixing.....	28
2.2.1. The Matrix Language Frame model and the markedness theory.....	34
2.2.2. Muysken's (2000) typology of code-mixing.....	38
2.3. Historical code-switching vs. lexical borrowing.....	43

III. The Genesis, Morphology, and Semantics of Place-Names

3.1. The appellative origin of place-names.....	52
3.2. Place-names as darkened compounds.....	58
3.3. Sound change and morphological change in place-names.....	63
3.4. The genesis of hybrid toponyms.....	67

IV. Layers of Language and Settlement in Britain

4.1. Defining the broad historical context.....	71
4.2. Pre-Celtic and Celtic Britain.....	72
4.3. Language in Roman Britain.....	75
4.4. Anglo-Saxon England.....	77
4.5. The Viking invasion of England.....	80
4.5.1. The extent and nature of the Viking settlement.....	81
4.5.2. Language in Viking Age Britain.....	84
4.6. The end of the Viking era and the Norman Conquest.....	87

V. Stratification of British Place-names and Hybrid Names

5.1. The corpus of settlement names.....	92
5.2. Methods of analysis.....	95
5.3. A note on pre-Celtic names.....	96
5.4. The Celtic layer.....	98
5.4.1. Celtic generics and specifics.....	102

5.4.2. Celtic hybrid names	103
5.5. The Latin element in English place-names	106
5.6. The Anglo-Saxon layer	108
5.6.1. Old English generics and specifics	110
5.7. The Norse layer	113
5.7.2. Scandinavian hybrid names	117
5.8. The Norman French layer	122

VI. Conclusions

6.1. Overview of results and concluding remarks.....	124
6.2. Possible directions for further research.....	127

VII. Magyar nyelvű összefoglaló

7.1. Kutatói kérdések és a disszertáció célja.....	131
7.2. A kutatás elméleti háttere.....	132
7.3. A korpusz bemutatása	134
7.4. Az empirikus kutatás eredményei, konklúziók és kitekintés	135

References

List of symbols and abbreviations

CS	code-switching
CM	code-mixing
CL	congruent lexicalization
ML	matrix language
EL	embedded language
MLF	Matrix Language Frame model
OE	Old English (c. 450 – c. 1100)
ME	Middle English (c. 1100 – c. 1500)
ModE	Modern English (c. 1500 – present day)
ON	Old Norse
NF	Norman French
CGmc	Continental Germanic
pn	place-name
pers.n	personal name
RO	rights and obligations
SSNY	Scandinavian Settlement Names in Yorkshire (Fellows-Jensen 1972)
SSNEM	Scandinavian Settlement Names in the East Midlands (Fellows-Jensen 1978)
SSNNW	Scandinavian Settlement Names in the North West (Fellows-Jensen 1985)
DB	Domesday Book (1086)
<	developed from
>	developed into
←	borrowed from
→	borrowed into
*	reconstructed form

To the memory of my father

I. Introduction

1.1. Introduction

The phenomena of code-switching (CS) and code-mixing (CM) have received considerable amount of scholarly attention, with quite a heavy focus on syntax, especially on two major syntactic types: intrasentential and intersentential switching, analyzed from a chiefly synchronic perspective (e.g. Poplack 1980, Myers-Scotton 1993a, Muysken 2000, 2011, Bullock & Toribio 2009a, Gardner-Chloros 2009), while historical aspects of CS have been somewhat backgrounded and still receive less attention (e.g. Schendl & Wright 2011, Schendl 2012). In contrast with this, this dissertation is concerned with diachronic code-mixing on a lexical level, as it is observable and contained within the boundaries of compound words that function as appellative place-names.

The aim is to investigate the nature of CM found in hybrid settlement names of Britain, with a pronounced focus on Scandinavian-influenced and Anglo-Scandinavian hybrid place-name formations in the Danelaw area. It is argued that these English-Norse hybrid names and the various manifestations of cognate substitution found in Scandinavianized settlement names are in fact instances of code-mixing consistent with Muysken's category of congruent lexicalization (Muysken 2000: 122-153). In the case of other language pairs (i.e. Celtic-English, Latin-English, Celtic-Latin, Celtic-Scandinavian) borrowing is postulated to be the main route of hybrid toponym formation. The reason why Scandinavian-related names are at the focal point of the investigation is that they are by far in the largest numbers amongst hybrid toponyms in England, and they are the ones that are most likely to have been created by a bilingual speech community. Altogether six research questions pertaining to historical code-mixing and borrowing in general and to the specific hybrid English toponyms will be formulated and answered in this dissertation (see Section 1.3 below).

The empirical analysis will be carried out on a corpus of British place-names, containing altogether 10,311 entries of which there are 924 hybrid names in total in the following breakdown by language pairs: 596 English-Scandinavian hybrid formations, 198 English-Celtic hybrid names, 3 English-Latin hybrids, and 17 English-Norman French hybrids. For a detailed description of the corpus itself, the sources of data, the method of data collection and analysis and the distributions of various hybrid place-name formations see Chapter V.

The present dissertation builds on and largely incorporates my previous, preliminary study (Fekete 2015) on English-Scandinavian hybrid place-names, with that paper and its accompanying corpus forming the nucleus of the dissertation's empirical analysis and the enlarged corpus compiled for it. This dissertation aims to provide a wider perspective of hybrid toponyms than just Anglo-Norse ones, while also expanding on the theoretical background of code-switching and code-mixing and including a discussion of the structural and semantic characteristics of place-names in general. I will also examine the various processes of language change that affect proper names.

Besides the corpus-based analysis of hybrid names created by the various layers of settlers to the British Isles, the dissertation also deals with historical linguistic aspects of the emergence of these toponyms. The place-names are treated as regular and originally transparent compound words, and are analyzed from a morphological viewpoint. The entire analysis is embedded in the framework of historical sociolinguistics and in the framework of code-mixing through Myers-Scotton's markedness theory (1982, 1989) and Muysken's (2000) taxonomy of code-mixing, with special attention to his treatment of congruent lexicalization. All of these theories and analytic frameworks were originally put forward with a synchronic perspective in mind, and in the present dissertation their tenets and principles will be applied for speech communities from a diachronic aspect.

1.2. Background and purpose of the dissertation

The purpose of this dissertation is to examine the phenomenon of code-switching and to a lesser extent borrowing in the various layers of English toponyms, paying special attention to Scandinavianized names and English-Scandinavian hybrid place-names. The starting point of the entire investigation can be summarized with Fellows-Jensen's (1980: 192) observation, that

“[b]etween about 600 B.C. and 1066 A.D. there were no less than five foreign conquests of England. Each of these conquests was accompanied by settlement and each wave of settlement resulted in both the coining of new place-names and the adaptation of old place-names to forms more congenial to the tongues of the newcomers.”

Therefore, the historical backdrop of the research is the fact that the English language and its speakers were engaged in direct linguistic contact situations with three different

languages (Celtic, Norse and Norman French) each some 300 years apart, and all of these contacts resulted in observable changes in and influences on the English language, including place-names. The contact of English with Latin is less direct than that of Celtic with Latin was, because Latin elements entered the English language primarily through Christianization and cultural prestige and not through conquest. The various waves of conquest and settlement that Fellows-Jensen (1980) quotation above talks about led to the emergence of contact situations which varied in closeness, depth, and extent. The first known settlers of the island who are relevant from the historical linguistic viewpoint of the dissertation were the Celts. *Figure 1* below represents a generally accepted schematic timeline of conquest events and the languages in use in those eras during the history of England. The gaps between the arrivals of the different groups in the figure serve the purpose of mere illustration, and are obviously disproportionate compared to the actual length of the time periods.

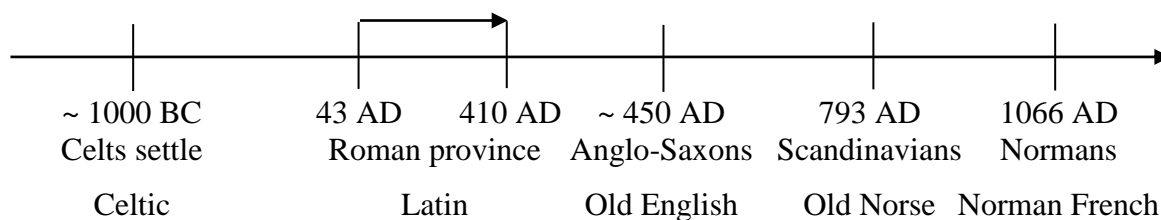


Figure 1. The timeline of conquests in England

This schematic representation gives a good impression about the extent of multilingualism that was present in medieval England, and it also reflects the external influences that English was exposed to during its history spanning more than one and a half millennia. Before the arrival of the Celts in the 1st millennium BC, several tribal groups mostly of Iberian origin inhabited the island, who will collectively be referred to as ‘pre-Celtic’ in future discussions. Little is known about this pre-Celtic population, although they did leave a mark on English place-names, mostly in the names of large geographical features, as a number of surviving place-names, river-names, and place-name elements testify to their existence. Questions of the pre-Celtic inhabitants of Britain will also be briefly discussed in the dissertation in the relevant subchapters.

The present dissertation also aims to examine hybrid settlement names embedded in a socio-historical context. The reason why names are in the focal point of this dissertation is twofold. One reason is that studies of code-switching and code-mixing, as pointed out above,

have been preoccupied with describing mostly synchronic syntactic phenomena and that the study of names has been a rather neglected area in linguistic analyses and tend to be the focus of scholarly attention from a chiefly onomastic and language philosophic point of view. As van Langendonck (2007: 2-3) also observes,

“[t]heoretical linguists have often treated proper names as the poor cousin of other grammatical categories. [...] Onomasticians, however, have sometimes forgotten that proper names are part of the system of natural languages. Both onomasticians and linguists should be aware of the fact that proper names are words which deserve linguistic attention in the first place.”

With this dissertation therefore, I wish to involve hybrid toponyms in the study of code-switching, language contact phenomena, and historical sociolinguistics, broadening the scope of these disciplines. The other reason is that hybrid place-names carry important pieces of evidence about the sociolinguistic environment in which they emerged, which can be accessed by studying the patterns and use of generics and specifics originating from different languages within one toponym.

Owing to the fact that place-names originate in language as semantically transparent lexical items, they will be treated as regular linguistic elements, with regular morphological and semantic characteristics (a detailed discussion of related issues will be presented in Chapter III). The main underlying assumption for their analysis will be that they behave similarly to non-proprial compound words which undergo formal and semantic obscuration, i.e. become darkened compounds during the historical development of language, and this process of obscuration is similar to the one that lexical compounds undergo.

The novelty of the present dissertation lies in the facts that (i) it deals with historical code-switching on the level of the lexeme as opposed to previous and mainstream approaches which are almost exclusively of a syntactic nature; (ii) it examines place-names from a morphological and phonological perspective in a socio-historical context and from a chiefly historical linguistic angle with special attention to language contact phenomena and processes of language change thereby fusing the description of changes observable on various levels of linguistic description with historical sociolinguistics; (iii) it does not consider proper names to be imbued with special qualities that would make them exempt from regular linguistic analysis; and (iv) it treats them as opaque compounds which were once semantically transparent and treats them in general as lexical items which behave like non-proprial lemmas

albeit they are more prone to certain structural changes and are more likely to preserve certain lexical elements which have become extinct in the language at large (cf. Coates 2006a).

Quite a vast amount of previous research is available on the central topics of the dissertation, however, to my knowledge at least, there has not been any investigation conducted on them in this specific way. Place-names have been mainly analyzed from a taxonomic and lexicographic perspective (e.g. Ekwall 1980, Mills 1998, 2011, SSNY, SSNEM, SSNNW), or with the purpose of providing a general overview and framework of analysis for them (e.g. Reaney 1987, Cameron 1996, Hoffmann 2007). Also, countless research papers have been published on the etymological analysis of individual names¹ (e.g. Fellows-Jensen (1987) on the name of *York*, just to mention one of the many). Linguistic analyses of toponyms have focused mostly on their semantic, pragmatic, and syntactic characteristics (e.g. Anderson 2004, Coates 2006a, van Langendonck 2007, Colman 2014) or various theoretical or language philosophical questions (e.g. Algeo 2010). All of the above mentioned works contribute significantly to the theoretical foundations of the onomastic part of my dissertation, while the dictionaries serve as the backbone of my corpus. With my research in this dissertation and its forerunner paper, I intend to bring together fields of interest which have rarely if ever been combined before (such as historical code-switching and onomastics) and to explore new directions of utilizing corpora in historical linguistic research.

Concerning the classic treatments of place-names, Reaney's (1987) and to a considerably greater extent Cameron's (1996) monographs provide the main frame for analyzing and contextualizing the place-names of my corpus. In my dissertation, I also relied in part on the Hungarian tradition of place-name analysis, especially Ditrói's (2017) description of toponymic systems and name-giving models (Ditrói 2017: 9-38), and Hoffmann's monograph (2007) which provides quite a detailed and thorough taxonomy of the possible types of place-name formations in Hungarian, and also offers insights into the development and genesis of place-names and the ways in which such names can actually be analyzed from a linguistic point of view. In the present research, I focus on the internal structure of the place-names and on the meanings of the constituents and I treat toponyms as genuine lexical items.

The sociolinguistic aspect of names has also been brought under scrutiny in the scholarly literature (e.g. Nicolaisen 1975, Fellows-Jensen 1990, 1991, Udolph 2012), and

¹ In some cases, such as Aybes & Yalden's (1995) study about the possible distribution of wolves and beavers in Britain, these research papers can be very specific and interdisciplinary.

especially salient for the dissertation are questions of superstratal and substratal influences (e.g. Bölcskei 2012). Hybrid place-names, however, have not yet been analyzed as results of historical code-switching, although historical CS is a very current and rapidly expanding area of sociolinguistics. The role of language contact in the emergence of such toponyms is also quite a neglected area of research, and the emergence of hybrid place-names and what they can reveal about the sociolinguistic context in which they were coined has not been given a lot of attention either.

The primary focal point of the research is appellative hybrid toponyms, and to a lesser degree those place-names which have a personal name as their specifics. The reason behind this is that it can hardly be construed as code-switching when a personal name from a different language is used in a recipient language, because with genuine CS speakers have a choice of which language's elements to use in their output, whereas with a foreign name no such choice can be made. Therefore, while such names will be considered hybrids in the dissertation, they will not be considered to be instances of code-switching. With this in mind though, personal name hybrids are still valuable sources of historical sociolinguistic data, and they will be given equal consideration and treatment in the dissertation as other types of hybrid names, but will not be construed as CS.

Summarizing the reason for choosing this topic, the dissertation deals with hybrid place-names because they reflect the interaction of speech communities therefore they can be studied from the perspective of historical sociolinguistics especially concerning the linguistic manifestation and consequences of substratal and superstratal influences. Furthermore, place-names have not yet been utilized to research aspects of historical code-switching, and they might prove to be useful tools for exploring linguistic borrowing as well.

This dissertation aims to address the above outlined issues and problems in a systematic manner, first surveying the already existing theories and available research results, then proceeding to analyze a corpus of English settlement names, and interpret the results in the light of the theoretical framework. The research questions of the dissertation will be formulated in the next section.

1.3. Research questions

The research questions of the dissertation can be divided into two different types. One half of the questions (Questions 1-3 below) are more of a theoretical nature, and they will be answered by the literature review and by the discussion of the theoretical underpinnings and frameworks of the empirical analysis outlined in Chapters II-IV. The other half of the questions (4-6) will be answered by the corpus-based empirical research in Chapter V. The answers to the questions formulated below will be given and discussed at the end of the dissertation, in Chapter VI. The dissertation is concerned with examining and answering the following research questions:

- (1) How is it possible to draw a demarcation line between historical code-switching and lexical borrowing?
 - (1a) How can historical code-switching be embedded in a synchronic theoretical framework of code-switching?
 - (1b) Is it even necessary to distinguish sharply between these two phenomena?

This first question is a theoretical one, aimed at examining the possibility and necessity of distinguishing historical CS and borrowing. The issue of distinguishing the two phenomena has been addressed in the literature from the perspective of synchronic code-switching (e.g. Myers-Scotton 1989, Backus & Dorleijn 2009) but the question itself is still yet to be answered. The purpose of this research question is to provide a synthesis of arguments from the literature pertaining to synchronic aspects of CS and CM, and give an answer concerning the diachronic dimension based partly on the results of the empirical analysis and partly on extrapolation from synchronic theories.

- (2) Are there certain levels of language (i.e. morphology and phonology) which are not or cannot be affected by code-switching?
 - (2a) How can phoneme substitution and cognate substitution (as described by Townend 2002) be accommodated in a model of code-switching?
 - (2b) How is phoneme substitution and cognate substitution manifested in hybrid place-names in relation to the sociolinguistic context in which they emerged?

This research question is related in part to the constraints that have been formulated for code-switching (e.g. Pfaff 1979, Poplack 1980, Myers-Scotton 1993b, Gardner-Chloros & Edwards 2004) and in part to the analysis of hybrid place-names which came about via cognate substitution and sound replacement, most of which are English-Scandinavian hybrids or Scandinavianized names. The second sub-question pertains to the corpus analysis, and aims to investigate whether any kind of correlation exists between the utilization of element substitution and the stratal relationship between the pair of languages.

- (3) How can the emergence of hybrid toponyms be analyzed in the framework of historical code-switching and historical sociolinguistics?

This question is related to the emerging field of historical code-switching (Schendl & Wright 2011) and will be answered by a synthesizing discussion of a cross-section of the available scholarly literature on the subject along with my own suggestions regarding the question on the basis of the corpus analysis. In order to get valid and consistent results, the notions of code-switching, code-mixing, and hybrid place-names need to have a clear definition, which I will provide in the discussion. Clear terms will also be defined as to what constitutes code-mixing in the case of hybrid toponyms, and what kinds of names are to be considered hybrid ones.

- (4) How is the process and the outcome of code-switching and place-name hybridization influenced by the genetic relationship and sociolinguistic status of the languages involved?

Moving on to the exclusively empirical set of research questions, the fourth one pertains mostly to the stratal relationship of the languages involved in the creation of the hybrid names. English is genetically related to all three of the languages that it came into direct contact with (i.e. Celtic, Norse, and Norman French), of which Old Norse was its closest relative. A degree of mutual intelligibility is likely to have existed between Old English and Old Norse which facilitated the transfer of loanwords and the creation of hybrid place-names. It will be argued that Old English and Old Norse were in an adstratal relationship, with superstratal tendencies of Old Norse. Concerning the other two languages, Old English was clearly superstratal to Celtic and substratal to Norman French. This question seeks to investigate to what extent these relationships are reflected in hybrid toponyms and

the usage of place-name elements. It is hypothesized that the adstratal status and mutual intelligibility were key contributing factors in the emergence of English-Scandinavian hybrids of every type, including ones created via element substitution.

- (5) How is historical code-switching manifested in hybrid place-names, and what patterns can be observed?

The final two research questions of the dissertation are the most important ones, as providing an answer for them is the main purpose of the entire corpus analysis. Question 5 deals with the actual analysis of hybrid toponyms, during which the various types of generics and specifics that occur in hybrid names will be examined and contrasted with their occurrence in monolingual toponyms. The way in which code-switching is observable in these names will be highlighted and the frequencies and distributional patterns of generics and specifics from different languages will also be provided and discussed. The examination of regional differences in the use of generics is also relevant for this question, therefore the notion of onomastic dialects (Nicolaisen 1980) will be of importance here, as different regional variants of the same language very often exhibit different naming patterns and preferences for using different place-name formants.

- (6) What kind of hybridization processes are observable in the case of English place-names, and what can these hybrid names reveal about code-mixing and the linguistic contact situations in which they emerged?

The other main question of the dissertation is related to the analysis of the empirical results in the light of the theoretical and socio-historical background outlined in the first half of the thesis to see how those theories correlate to actual place-name data and to examine the various routes through which hybrid formations can come about. The nature of contact situations that English entered into with different languages will also be correlated with the hybrid place-names to see what patterns and tendencies are observable in the case of English-Celtic, English-Scandinavian, and English-French interactions. The context and nature of the contact situation is hypothesized to be determinative of the outcome of hybridization and the utilization of code-switching or borrowing. The possible origin, i.e. code-mixing, borrowing, or element substitution, of different types of hybrid toponyms will also be examined.

1.4. Structure of the dissertation

The dissertation is divided into eight chapters, the first of which serves the purpose of an introduction and the last contains the final conclusions and a recapitulation of the results of the research. The second and third chapters provide a theoretical background for the investigations presented in the dissertation, with Chapter II describing and providing a critical overview and evaluation of the major theories of code-switching and contact linguistics relevant for the emergence and analysis of hybrid settlement names, while Chapter III discusses in a similar manner those theories of morphology, semantics, language change and onomastics that are relevant for the research, such as compounding, the appellative origins and denotative function of place-names, the role of folk etymology and reanalysis and semantic transparency. Chapter IV focuses on providing a more detailed historical background from pre-Celtic populations until the arrival of the Anglo-Saxons. Beside giving an overview of historical events, this chapter also discusses the linguistic situation and the sociolinguistic relations of the languages spoken during various eras of conquest as well as the settlement areas of the population groups involved. The fifth chapter deals with the Scandinavian invasion of England and the subsequent settlement of the Vikings, also taking into account their linguistic influence and issues of bilingualism and mutual intelligibility between the Norse invaders/settlers and the Anglo-Saxon population. The next chapter is centered around examining the various layers of toponyms found in Britain, describing, first of all, the corpus of settlement names in detail, then discussing and analyzing the distribution of Celtic, Roman, Anglo-Saxon, Scandinavian and French generics and specifics, as well as the various hybrid toponym formations which emerged through the contact scenarios these languages entered into with each other. In Chapter VII, the data of the previous chapter is interpreted in the light of the theories introduced in Chapter II and III, and the research questions in Section 1.3. are answered.

Finally, the dissertation concludes with a general summary and concluding remarks, followed by a separate chapter containing the Hungarian recapitulation and summary of the dissertation's main research questions, materials, and empirical results. The dissertation ends with the list of references and 3 appendices which contain the full list of the generics and specifics that occur in the corpus, in a language-by-language breakdown, complete with meanings, frequency counts and percentages, and regional distributions.

1.5. Limitations of the research

Naturally, every research has its limitations and shortcomings and this dissertation is no exception either. Perhaps the biggest challenge I faced while writing it was that the size of the overall existing place-name material is far larger than those **13,000** instances that form part of this investigation. Only those names are featured in this dissertation and the accompanying corpus which could be found in the sources that were available at the author's disposal (which are the following: Ekwall 1980, Mills 1998, 2011, SPLNY, SSNY, SSNEM, SSNNW) which include only a fraction of all the possible English place-names.

The English Place-Name Society's (hitherto) 91-volume county-by-county survey record, which is the richest and most thorough collection of English toponyms, was sadly not available to be used as the basis of my corpus. Perhaps the greatest advantage of the EPNS survey is that it contains information about lost place-names, something that is absent from Mills' (1998, 2011) and Ekwall's (1980) dictionary, as these focus only on those place-names that have survived until the present day and unfortunately they do not have any records about lost names and depopulated settlements. Still, fortunately, Fellows-Jensen's (1972, 1978a, 1985) surveys do provide data about lost toponyms and those are included in the corpus. The EPNS surveys have a wealth of information available (such as for instance lost names, many more attested forms than in Ekwall's and Mills' dictionaries and Fellows-Jensen's monographs, microtoponyms, field-names, street names, etc.), and ideally a corpus should be constructed from those publications, but it is quite probably humanly impossible for one person to sift through and encode all of that by himself, especially in the limited time that was available for the completion of the present dissertation. It would require an entire dedicated research team and years of work to cover all of it. The sheer size of the EPNS project is not only indicated by the 91 volumes published since the 1924 launch of the series but also by the fact that it is still incomplete and there are entire counties in England which have not been surveyed yet.

The conclusions drawn here are therefore to be understood in relation to the toponomastic material that is collected and included in the corpus, which, however extensive it may be, still cannot cover the entirety of the English place-name material neither in breadth nor in depth. Bearing this in mind, the dissertation still offers an interdisciplinary approach to language contact, historical linguistics, and toponomastics, and aims to provide a solid basis for future research in these areas.

II. Code-switching and other Theoretical Considerations

In this chapter those theories and frameworks of analyzing code-switching from either a synchronic or a diachronic perspective will be highlighted and surveyed which bear relevance to the formation of hybrid place-names in English. In particular Muysken's (2000) typology of code-mixing will be important, especially his category of congruent lexicalization (Muysken 2000: 122-153) and to a lesser degree that of insertion (ibid 60-95). Congruent lexicalization is a type of code-mixing whereby elements of the two languages involved are inserted into structures that are shared by both of the languages (Muysken 2000: 127), while insertion refers to that type of code-mixing whereby constituents of a language are embedded into the structure of another language. Congruent lexicalization primarily occurs between closely related languages (very often between varieties of one language: dialects, or a dialect and the standard variety) that share a common grammatical structure. Insertional code-mixing on the other hand is often analyzed in the context of lexical borrowing, and is seen by many researchers to be part of a continuum that also encompasses transfer and borrowing. Features of these two categories and Muysken's third type of CM, alternation, will be discussed in this chapter, along with their relevance for the emergence of hybrid toponyms. The notions of CS and CM can be defined in a variety of ways, as a preliminary stance, I will take CS to refer broadly to switching occurring in between sentences, and CM to refer to switching occurring within sentences.

Myers-Scotton's Matrix Language-Frame model (1993a) and markedness model (1983, 1989) will also be discussed in a separate subchapter and it will be demonstrated that in the context of English-Scandinavian hybrid place-names the distinction of a matrix language and an embedded language is not perfectly justified because congruent lexicalization obtained in those cases. The number of hybrids with English generics and Norse specifics is roughly the same ($N = 254$) as the number of place-names that have ON generics and OE specifics ($N = 342$), therefore it cannot be established with absolute certainty whether OE elements were inserted into an ON frame or vice versa. In the case of other language pairs English acted as the matrix language, and Celtic, Latin, and Norman French as the embedded language.

Attention will also be devoted to more general theoretical questions, such as the notion of code-switching and code-mixing, their possible differentiation, and some of the numerous constraints that have been proposed for code-switching, all of which are ultimately flawed, as evidenced by the countless counterexamples that violate these constraints. The main focus

will be on Poplack's (1980) free morpheme constraint and equivalence constraint, with a discussion of their relevance in the case of hybrid toponyms, and a survey of counterarguments from the literature and counterexamples from the corpus analysis will also be provided.

Since the notion of borrowing is also very important in the discussion of code-switching, as the two phenomena are strongly tied to each other, and can be conceptualized as part of a continuum with fuzzy boundaries. The final subchapter of this part of the dissertation will be devoted to the discussion of the correlations and differences between CS and borrowing. In the discussion, the term borrowing itself will be clarified, and compared and contrasted with historical code-switching. The phenomenon of borrowing will be discussed in the framework of the markedness theory described by Myers-Scotton (1983, 1989) and Muysken's (2000) code-mixing typology, concluding with Backus' (2015) usage-based model which incorporates CS and borrowing into a unified model. Finally, the relevance of borrowing for hybrid place-names will also be discussed.

2.1. The notion of code-mixing and its relevance for hybrid place-names

Before proceeding to the discussion about the nature of code-switching and code-mixing, first and foremost giving a definition of hybrid place-names shall be in order. In the broad sense, a hybrid place-name is a toponym which consists of at least two elements, one of which comes from a language that is different from the language of the other element in the toponym. In a narrow sense, however, only those place-names can be considered to be genuinely hybrid ones which did not come about through borrowing name formants from another language and utilizing them to create new names in the receiving language, but ones which were (likely) coined and accepted by a bilingual community via utilizing specifics and generics from both languages without the involvement of any borrowing. While the usage of both the broad and the narrow definitions can be justified, only the narrow one entails and permits code-mixing.

In his pioneering treatment of language contact, Weinreich (1953) provides a discussion of lexical interference in which he touches upon hybrid compound words and place-names too (Weinreich 1953: 50-53). In his account, he notes that lexical interference can be manifested in compounds and toponyms in the form of transfer of elements or reproduction of elements. This means that a compound can either be fully transferred from the donor language of only one element is taken over while the other is supplied from the word-

stock of the recipient language. Weinreich also considers those words to be hybrid compounds in which the stem of a complex word is transferred and the derivational affix comes from the receiving language, which in modern analyses would not be construed of as a compound. In the case of hybrid toponyms, Weinreich points out that analyzable place-names are sometimes translated into the recipient language element-by-element, and in hybrid names only one element is transferred and the other is substituted for by a native element.

Gammeltoft (2007: 481) dismisses the claim that hybrid names exist, arguing that all seemingly hybrid formations are in fact monolingual ones which emerged as a result of borrowing place-name elements which were used productively in the receiving language (see the discussion in Sections 5.6 and 5.7 below about English-Scandinavian names for further details). Walther (1980: 144) on the other hand seems to be more permissive, yet he also formulates a restriction about what is to be construed as hybridization: in his view this term can refer only to the partial integration and adaptation of multi-element names, and the complete adaptation cannot be labeled as hybridization. This is a sound and logical point because if a foreign name is completely taken over and gets integrated into a language in its entirety, then that name can no longer be seen as a foreign element. If a name is only partially taken over by a language then one of its elements will remain foreign, therefore a hybrid name emerges. This criterion, however, does not take into account those names which are created originally as hybrid names by a bilingual society, as it is concerned only with hybridization via borrowing and integration. The broad and narrow definitions I put forward above are still in agreement though with Walther's criteria for hybridization.

In a discussion of the possibility of medieval Slavic-German individual bilingualism in the territory of present-day Austria on the basis of Slavic-Germanic hybrid place-name formations in the area, Holzer (2015) argues that loan translations and hybrid name systems can be taken to be indicators of individual bilingualism. In his analysis, Holzer differentiates between bilingualism on a geographical level and bilingualism on the individual level, with the former referring to a setup whereby a territory is shared by different groups of people each speaking their own language separately, and the latter referring to an individual's ability to speak two languages. Territorial bilingualism is sufficient for place-names to be borrowed, therefore borrowed names and "etymologically bilingual" (Holzer 2015: 9) names alone do not necessarily indicate that the individual inhabitants were also bilinguals. The conclusion is that hybrid place-names can only be indicative of individual bilingualism if they are of a specific type, namely calqued toponyms, given that in order to translate a name element one has to be able to understand that language. Other types of hybrid names could have been

formed by analogy after one of the languages had ceased to be spoken in the area. In the case of toponyms in my corpus, Celtic-English hybrids likely arose through the borrowing and extension of already existing place-names upon the arrival of the Germanic tribes, while English-Scandinavian hybrids possibly came about through cognate substitution (these would correspond to Holzer's calqued names) and as originally hybrid names.

Apart from the above cited passage from Weinreich (1953: 50-53) and papers by Walther (1980), Gammeltoft (2007), and Holzer (2015), of which only Weinreich's and Gammeltoft's contribution is written in English, to my knowledge the scholarly literature is rather silent about the topic of hybrid place-names, save for a few isolated, passing mentions of the phenomenon. Even fewer sources can be found which deal with the possible relationship of hybrid toponyms and code-mixing. Sköld (1980) could perhaps be considered to somewhat belong among them, albeit he works with quite a different understanding of CS and his discussion is not specifically about hybrid names, but the integration of foreign names into a recipient language.

In a more recent publication on the subject, Martynenko (2015: 12) offers a brief definition of hybrid toponyms, describing them as "place names composed of lexical and/or grammatical means of two languages" which in itself is in accordance with the possible definitions I introduced at the beginning of this section. In her brief paper Martynenko deals with English-Spanish hybrid toponyms in the United States, but her focus is more on the historical aspects of the emergence of such place-names and giving a taxonomy and extensive lists of them, and less so on the actual linguistic analysis of the names themselves. Cameron (1996) also mentions hybrid names, but it is again restricted to taxonomizing the names and exploring the socio-historical context in which they emerged. Minkova & Stockwell (2009: 36) treat Celtic place-names such as *Kent*, or *Dover* and place-name elements like *-combe*, and *-torr* in English as borrowings and list the names *Yorkshire*, *Devonshire*, and *Canterbury* as hybrid Celtic-English toponyms (cf. Cameron 1996: 57, 60). In their discussion of Scandinavian loanwords, they consider the elements *by*, *thwait(e)* and *thorp(e)* to be borrowed elements in English (Minkova & Stockwell 2009: 39). This interpretation of these Celtic names is valid, as the simplex names and the generics were likely borrowed, and the multi-element names are hybrids in the broad sense, because they contain a foreign element, although they were probably not coined by a bilingual speech community. Concerning code-switching and proper names, Baukó (2015: 100) makes a very brief, passing reference as part of his discussion on contact onomastics, mentioning only that proper name code-switches can be encountered in written and spoken language, and that during code-switching more than one

language is given an active role within a single discourse. In his monograph, Baukó (2015: 103) uses the term ‘code-switching’ to refer to Hungarian and Slovak personal names being used in Slovak and Hungarian discourse, respectively.

In theory, the primary focus of this dissertation should be those hybrid names which conform to the narrow sense of my definition, but we must bear in mind that in reality for most of the time, it is quite difficult to tell apart which names are genuine hybrids and which ones involve borrowing. Still, we can at least to some extent rely on the frequency of occurrence of given specifics and generics to arrive at an approximate conclusion as to the history and origin of a given name. A high frequency of occurrence would likely hint at borrowing, while a low frequency would indicate that the element in question was used on a one-off basis characteristic of code-mixing behavior. The diversity and type/token ratio of generics that occur in these names can also be indicative of whether they arose through CS or borrowing. If there is only a limited amount of different generics that occur frequently then the case is likely to be borrowing, while if there is a wide variety of generics, many of which occur with low frequencies, then code-switching could prove to be the more plausible explanation. These frequencies, ratios, and their implications will be discussed in greater detail in Chapter VI.

For the purposes of this dissertation, the amalgamation of two broad definitions of code-switching (and code-mixing) is adopted, namely that code-switching is “the alternation of two languages within a single discourse, sentence or *constituent*” (Poplack 1980: 583 [emphasis mine]) and also in an even broader approach it is “the use of more than one language during a single communicative event” (Muysken 2011: 301-302, cf. also Poplack 2015: 918). In analyses of code-switching this “communicative event” is usually taken to be a sentence or an utterance, yet in our case it will be further narrowed down to the production of a compound word, which is then used as a place-name, hence the adoption of Poplack’s (1980) definition which includes code-switching within constituents.

Myers-Scotton (1989: 336) adds to this two slight, yet very important modifications, namely that CS is the “*speaker-motivated* use of two or more *linguistic varieties* (language, dialects or styles) in the same conversation” (emphasis mine). The claim that code-switching is speaker-motivated rests on the valid assumption that it is ultimately the speaker who is in control of their linguistic output. However, external influences can interfere with language use and dissuade the speaker from making their original choice. Therefore, strictly speaking, any instance of code-switching that is necessitated by speaker-external forces, such as a change of topic or the person of the interlocutor, is not to be considered genuine CS under this

definition. This is problematic because this approach would discard many actual code-switches on the grounds of them not being the result of the speaker's conscious decision to switch. This is an unnecessary restriction, especially compared to the more result-oriented approach that I adopt, whereby any concurrent use of segment from two different languages is considered CS. In this respect, I agree more with Pfaff (1979) who refers to code-switching as a socially motivated phenomenon. I should note here, however, that in this dissertation I do not consider those hybrid place-names that have a personal name as their specific to be true instances of CS, because (as outlined in Section 1.2) foreign personal names are typically used in their original form or rendered as a close approximation therefore the speaker has no other choice but to use those names the way they are (cf. also Holzer 2015: 9 about the use of personal names in hybrid toponyms).

Myers-Scotton's other extension of the definition involves the broadening of the possible scope of CS by including switching between dialects or styles and not just separate languages. This is a very important addition, as it covers a host of different contexts and situations of code-switching, even including what would otherwise be subsumed under the label of style shifting and traditionally not considered to be code-switching.

In her highly influential paper, Poplack (1980) recapitulates and elaborates on two major constraints that she developed earlier and that still have currency in today's CS research, and should hold true for all instances of CS: the free morpheme constraint and the equivalence constraint. According to the free morpheme constraint, codes can be switched "after any constituent [...] provided that constituent is not a bound morpheme" (Poplack 1980: 585), which means that no switches should occur between two inflectional or derivational morphemes and between a free and a bound morpheme. The equivalence constraint holds that switches are likely to occur at those points where the two languages share their structure or have some sort of structural similarity or congruity, which means that syntactic convergence facilitates the process of code-switching. As Poplack (1980: 586) describes it "a switch is inhibited from occurring within a constituent generated by a rule from one language which is not shared by the other."

These constraints are relevant in our case, given that hybrid place-names will be analyzed from the perspective of a specific type of code-switching known as congruent lexicalization in which, as will be seen from the discussion below, structural similarity is a key notion and requirement (albeit this similarity is mostly taken to be in the syntactic structure of the languages). The free morpheme constraint is also important in the case of place-names because switching and hybridization in these formations tends to occur between

free morphemes (i.e. the constituents of the place-name). However, those English-Scandinavian hybrid names which contain a Scandinavian inflectional affix in their specific seem to violate the free morpheme constraint because they contain an English word in the generic after the Scandinavian inflectional bound morpheme found in the specific, a position where, as per the constraint, they should not occur at all. Due to the fact that for the emergence of English-Scandinavian hybrid toponyms a shared linguistic structure was a prerequisite, those names do not violate the equivalence constraint, and they also largely conform to the free morpheme constraint due to the reasons outlined above. Poplack's constraints will be taken up again and surveyed in greater detail in the next subchapter.

The practices of code-switching and code-mixing are closely related to bilingualism, and engaging in them requires a degree of bilingual competence (cf. Muysken 2011: 304). If speakers engage in the process of code-switching or code-mixing, they possess a degree of linguistic competence in both of the languages they mix, and in case the two languages involved are genetically related to each other, then essentially the speakers perform an act of cognate substitution. The term 'cognate substitution' itself refers to the process whereby speakers of two related languages, ideally closely related ones, engage in the act of code-switching during which the speakers substitute a word in their language with the cognate of that word from the other language. Due to the fact that the two languages are close relatives of each other, the speakers are essentially swapping in cognate words during communication. The message and the communicative content that the speaker intends to convey this way is exactly the same as it would be if it was conveyed in only one of the languages and the only aspect that is different is the surface form².

In the case of hybrid place-names studied in this dissertation, bilingualism can be posited to have existed between Old English and Old Norse, but not between English and Norman French, or to a considerably lesser degree at least, as the majority of the French-speaking nobility did not learn to speak English. Regarding the Anglo-Scandinavian contact situation the bilingual competence was likely achieved through the existence of some mutual intelligibility (cf. Townend 2002). Drawing parallels with research results of analyzing the code-switching and language shift patterns of bilingual speakers who learned their second language in their adulthood, Ringe & Eska (2013: 74-75) argue that initially, speakers of Old Norse learnt Old English as adults and that Scandinavian speakers had shifted to the use of OE by the time the English reconquest of the Danelaw was complete in the mid-10th century.

² Code choice can also be the manifestation of one's conscious and deliberate expression of identity and can naturally bring about pragmatic issues.

Concerning the English-Celtic situation, it is also believed that bilingualism could have existed (cf. Schrijver 2007) which facilitated the transfer of Celtic river-names and place-names into English (cf. Poussa 1991: 300).

The phenomenon of code-switching is also strongly tied to the notion of borrowing, and the separation of the two is not always easy, especially so in the case of hybrid place-names. Apart from likely code-switches, a number of borrowed elements can also be found in hybrid place-names, as evidenced by the Celtic elements cited previously in this section, or by the very plausible borrowing of ON *bý* and *þorp* into Old English. In their seminal work, Thomason & Kaufman provide their definition of borrowing, namely that it is “the incorporation of foreign features into a group’s native language [while] the native language is maintained, but is changed by the addition of the incorporated features” (1991: 37), while in one of the earliest treatments of the issue Haugen (1950: 212) defined the phenomenon as “the attempted reproduction in one language of patterns previously found in another”. It is rather difficult to find the demarcation line between borrowing and diachronic code-switching, but as the previously cited definitions show, the key emphasis is on the concepts of *incorporation* and *reproduction*. Code-switching also makes extensive use of these processes. Mixed language utterances and lexical items show the incorporation of elements of one language in the structure of another, which are very often transformed both phonologically – so as to be more congruent with the phonological system of the recipient language – and morphologically through the addition of inflectional and other bound morphemes, thereby the mixed elements are reproduced utilizing the phonological and morphological inventory of the language into which the foreign elements are inserted.

2.2. The theoretical framework of code-switching and code-mixing

Before proceeding to the discussion of possible theories of code-switching, we shall first look very briefly at the field of CS itself. Backus (2015) identifies a number of main general strands of code-switching research, which are often thought to be rivals in binary opposition with each other. In terms of focusing on form or function we can have structural approaches which are concerned with the grammatical form of CS and with what principles constrain switching, and on the other hand the functional or sociolinguistic approaches are interested in how CS is used and what kind of communicative and pragmatic functions does it have. Concerning the functional perspectives, we have alternational code-switching and insertional code-switching, with the former concentrating on how and why do speakers use

chunks of two languages in an interleaved, alternating fashion, and the latter investigating why speakers insert individual, discrete lexical items of one language into the structure of a different language. The formal approaches tend to have a focus on how exactly are the two language systems used concurrently, and finally the psycholinguistic studies concentrate on the aspects of speech production in the case of mixed-language utterances.

In the literature on the subject, a difference is often made between code-switching and code-mixing, with the boundary between the two being drawn in various ways (cf. Gardner-Chloros 2009: 12-13). Generally speaking, in narrow approaches to the phenomena CS is taken to refer to switching of an intersentential nature (Kolehmainen & Skaffari 2016: 124-125), that is switches occurring between sentences, and CM is taken to refer to intrasentential switching, which occurs within sentences, arguing that CS and CM would require different cognitive processes, and different levels of integration of foreign rules into the matrix language (cf. Myers-Scotton 1989: 334, reviewing various approaches to CS).

Other sources focus on whether the codes in question show convergence towards each other, which would be characteristic of code-mixing, or they retain their individual features, which would be considered code-switching. For instance Muysken (2000) sets up the three-way distinction of mixing phenomena of alternation, insertion, and congruent lexicalization, which will be discussed in more detail in Section 2.2.2 below, depending on what kind of structure the foreign element is inserted into. He considers only alternation to be code-switching and insertion and congruent lexicalization to be code-mixing, as alternation means the alternating use of segments from each language and in this form it is closest to intersentential CS, and in insertion and CL foreign segments are inserted into the recipient language's structure or a shared structure, hence their resemblance to intrasentential mixing.

Contrary to this, Clyne (1987: 740-741) proposes quite a different configuration of notions. He considers code-switching to be “the alternative use of two languages either within a sentence or between sentences” meaning that he treats both intersentential and intrasentential switching as CS. Apart from providing a broad definition of code-switching, Clyne (*ibid.*) also introduces his term of ‘transference’ which means that a “single item is transferred from language B to A [...] whether integrated into the grammatical and/or phonological system of the recipient language or not.” This term then encompasses what traditionally would be equated with borrowing. However, from Clyne's definitions and discussion it emerges that ‘transference’ could cover both the notion of borrowing and of code-mixing. This is apparent from the treatment of CS as the “alternative use” of languages which could also hold true for code-mixing if we considered intrasentential switching to be

CM, but the real difference lies in the notion of transference. On the one hand the key phrase for defining CS is *use of languages*, which refers to two languages viewed holistically as a system being utilized simultaneously for expressing the speaker's communicative intention. On the other hand the key point of the definition of transference is *transfer of a single item*, which means that CM could be considered a form of transfer whereby only a single word is taken over that is not integrated into the recipient language's system. Like most researchers of code-switching, Clyne does not include in his definition switches which occur on the level of the lexeme, but he does bring up examples of German-English and Dutch-English switches within a single word and within bilingual compounds (Clyne 1987: 756). The occurrence of the latter is crucial for us, because switching within a compound is what happens in the case of hybrid place-names.

As it can be seen, the scholarly discourse is still to a considerable extent lacking a precise definition about what to be understood as code-switching and what as code-mixing. Due in part to this uncertainty of clear terms, and in part to the fact that CS and CM can occur concurrently while serving the same social function and discourse marking function, I echo and adopt Myers-Scotton's stance (1989: 334) that the distinction between CS and CM is "poorly motivated." Instead of trying to fit two separate labels to two similar surface manifestations of one underlying phenomenon I choose not to distinguish sharply between CS and CM in this dissertation, while still accepting and maintaining that differences can and do exist between the two notions. I take these differences to be the distinction between the interstential versus intrasentential nature of switching. In keeping with Muysken's (2000) analysis, I will be using the terms in an interchangeable fashion, although due to the fact that intrasentential switching has long been equated with CM, I would consider place-name hybridization to be code-mixing rather than CS. The rationale behind this is simply that in the case of hybrid toponyms the switching occurs within one lexical item, therefore it is more akin to intrasentential processes.

It has also been proposed in the literature (e.g. Sankoff & Poplack 1981) that speakers who engage in code-switching possess a code-switching grammar, or mixed grammar composed from the combination of monolingual grammars to generate utterances with CS, instead of having a separate grammar for each language. Such theories are grounded in the generativist tradition and assume that every native speaker's grammar is the same, and can be described as a set of standard norms. Based on her earlier analysis of English-Spanish intrasentential mixing Pfaff (1979) concludes that speakers engaging in CS possess grammatical competence in both of their languages and refutes the claim that a third grammar

would exist that speakers utilize when producing mixed utterances. Instead, the grammars of the languages are “meshed” according to certain constraints (Pfaff 1979: 314).

As it was outlined in the previous section, certain constraints have been proposed concerning the nature of code-switching, the most influential of which were Poplack’s (1980). It should be discussed here that both of her two constraints have received criticism, with perhaps the free morpheme rule being more susceptible to falsification (e.g. Clyne 1987: 758, Azuma 1996), although counterexamples for the equivalence of syntactic structures being a prerequisite have also been described (e.g. Clyne 1987: 757). The data of the present corpus, while conforming to the latter, also seems to run contrary to the former constraint because of the appearance of inflectional endings in the foreign specific before the generic, and because one toponym can contain elements from two different languages thereby forming a bilingual compound. In later papers Poplack herself revised and reformulated to some extent these theories (e.g. Sankoff & Poplack 1981), calling for their cross-linguistic examination in order to ascertain their universality. The free morpheme constraint has been modified to allow switching in positions that are adjacent to bound morphemes, provided that the lexical item in question is “phonologically integrated into the language of the bound morpheme” (Sankoff & Poplack 1981: 5), and the equivalence constraint (*ibid.*) is further clarified as such: the constituents in the immediate vicinity of the switch point “must be grammatical with respect to both languages involved simultaneously.”

Azuma (1996) criticizes and falsifies the free morpheme constraint by citing numerous cross-linguistic counterexamples against its operation, especially in the context of switching between English and agglutinative languages (cf. also Muysken 2000: 14). After identifying that the main pitfall of the free morpheme constraint is that it focuses more on the formal characteristics of the morphemes instead of their meaning, Azuma (1996: 366-367) proposes an alternative approach in which semantic content is key. According to the alternative stand alone principle, “any segment that can meaningfully stand alone in the speaker’s mind [i.e. a chunk] may be code-switched” (*ibid.*). Psycholinguistic research is cited which corroborates the theory that speech perception is carried out in chunks, and in speech production meaning comes first, and the attachment of linguistic form is secondary. This proposal is a significant improvement over the original free morpheme constraint, and it also has more positive implications for the hybridization of place-names, as it permits the switching of constituents (i.e. specifics and generics) as long as they form a meaningful unit in the speaker’s mental lexicon. The original constraint would have ruled out many actually attested hybrid names on the basis of them being impermissible. Given that even in isolation productive place-name

elements are lexically meaningful words, whether inflected, derived, or word-formation-wise left intact, they can be represented as meaningful chunks in the minds of the speakers.

Clyne (1987: 757-758) also provides counterexamples to the free morpheme constraint in the form of word-internal switches and hybrid compounds, and while acknowledging the validity of the equivalence constraint, falsifies the assumption that the syntactic systems of the languages in contact are stable, standard systems. Based on the discussion in this section, I maintain that the free morpheme constraint is overly grounded in the examination of the synchronic state of a relatively small set of languages, especially the largely analytic English, which also makes it Anglocentric.

Pfaff (1979), while also acknowledging that code-switching and code-mixing are subject to certain constraints, considers them to be socially and not structurally motivated. In her treatment of the phenomena she tangentially touches upon the diachronic dimension of language-mixing as well, which is quite a rare occurrence in the literature, albeit the questions she asks are not strictly related to historical code-switching, but are more of a theoretical nature³. One of the most important upsides of Pfaff's contribution is that she considers CS and language-mixing to be multi-faceted, i.e. that on the one hand they are motivated by social factors, and on the other hand they are subject to structural and usage-related constraints (cf. also Poplack 1980, Sankoff & Poplack 1981) and, perhaps more importantly, constraints that reflect the speaker's communicative intention and the semantic content of the utterance in which switching takes place (cf. also Azuma 1996).

Despite the vast amount of research into constraints of CS, attempts made at defining grammatical constraints and rules of code-switching, code-mixing and contact induced language change are ultimately largely unsuccessful at providing universal rules to which all instances of CS can and will conform (cf. Gardner-Chloros & Edwards 2004). While the constraints generally hold true and can be proven to be valid, most of them can also be falsified by adducing counterexamples (see Clyne 1987 for more constraints and examples, and see also Appel & Muysken 2005: 121-128 about constraints in general) which renders them questionable concerning their usefulness, validity, and reliability. Overall tendencies and principles of CS are to be observed instead of hard and fast rules. The linguistic behavior and code choice of the speaker greatly depends on pragmatic factors (such as power distance,

³ The questions related to language change that Pfaff (1979: 291) poses are: (i) how are different types of language-mixing (borrowing, code-switching) interrelated, (ii) is there any convergence amongst them, and (iii) does language change in multilingual communities parallel the development of pidgins and creoles. With the exception of the first problem, this dissertation does not aim to deal with such issues, therefore they will not be elaborated on further.

perceived prestige of the affected languages, familiarity of the interlocutors, context in which the exchange takes place, etc., cf. the discussion of Myers-Scotton's (1983, 1989) markedness theory in 2.2.1 below) and therefore it is difficult to lay down general rules. Instead of attempting to establish binary rules, it is perhaps more advisable to for instance observe and examine frequencies with which given types of switching are likely to occur in given types of communicative settings.

Moving away from the terminological quandary and the issues of structural constraints of code-switching, let us now turn our attention to the phenomenon as a larger picture. In an attempt to provide a unified approach to CS, Backus (2015: 20) calls for a "lumping approach" instead of fragmenting the discipline, and rightfully criticizes previous treatments of CS for compartmentalizing linguistic behavior and code-switching into separate sub-disciplines, which, I add, are often thought to be irreconcilable with each other (for instance Muysken 2011 also calls for a demarcation of CS from other contact phenomena). Such a perspective would also integrate synchronic and diachronic aspects of language contact, and would be able to account for the differentiation between code-switching and borrowing. Backus proposes a usage-based approach, which is built on a cognitivist foundation. The integration of cognitive linguistics, structuralist research, and most importantly social aspects of CS enables researchers to investigate the question in an interdisciplinary manner, something that I also aim to carry out in this dissertation.

Concerning code-switching from a diachronic perspective, by which I mean CS in historical texts and sources, Schendl & Wright (2011: 22-28) point out that historical code-switching is an exclusively written phenomenon. Code-switching in the Old English period was predominantly of an intersentential type and is restricted to non-royal charters and a few macaronic poems, while in the Middle English era, with the civic documents now produced in Latin instead of Old English and with the introduction of Norman French, the entire phenomenon changed quantitatively as well as qualitatively, although the medium still remained written. Due to the expansion that CS underwent in the trilingual Middle English era and the rise of macaronic poetry and sermons, and despite the relative scarcity of its occurrence in Old English, code-switching can be considered a standard feature of medieval multilingualism (cf. also Schendl 2012). Until the emergence of standard varieties and their codification and the rise of nationalism which also brought with it the issues of official languages, countries, states, nationalities and minorities in the era of Romanticism, societal multilingualism was quite widespread, so much so that it could even be considered the 'norm'

of the time, as exemplified by the history of English from the mid-5th century to the 14th century.

Finally, the last concept to be surveyed very briefly in this section is the relationship between code-switching and bilingualism. The term bilingualism itself can be defined in a variety of ways, and with very nuanced distinctions. A broad, pre-theoretical definition would include that a bilingual is someone who has “native-like control of two (or more) languages” (Bullock & Toribio 2009a: 7), which would entail that they have equal command of their languages, but in reality balanced or symmetrical bilingualism is extremely rare, if even existent. Rather, the extent of bilingualism is dependent on the context of language acquisition or learning, the age of learning the language, the (perceived) status of the languages, etc. It is a widely held belief, though, that in order to engage in CS the speaker has to be in possession of some bilingual competence (cf. Muysken 2011: 304), and that code-switching is an additional tool of expression for bilinguals (Bullock & Toribio 2009a: 8). In the context of the emergence of hybrid place-names bilingualism can be posited to have existed, which was not uniform in the case of every speaker. For our purposes and for the analysis of the context in which hybrid toponyms could have come about Myers-Scotton’s (2006: 44) rather broad definition of bilingualism will be accepted, namely that it is “the ability to use two or more languages sufficiently to carry on a limited casual conversation.” This approach and thus is in accordance with the dissertation’s theoretical basis as it rules out passive knowledge of a language as bilingualism and also rules out those speakers as bilinguals who only use borrowed foreign words but do not actually speak the language in question.

In the following three subchapters three important models of language change and code-switching will be outlined briefly that bear relevance to the theoretical framework of the dissertation.

2.2.1. The Matrix Language-Frame model and the markedness theory

Myers-Scotton’s markedness theory (1983, 1989) focuses on the so-called rights and obligations (RO) of speakers when engaging in communication and when making code choices. In essence, the purpose of conversation is to negotiate a set of rights and obligations, which derive from the “social features which [the speaker and the addressee] reciprocally accept as salient for the current exchange” (Myers-Scotton 1983: 117). The markedness theory also encompasses the rules and maxims that each participant must observe in order to ensure that their contribution is acceptable and normative in terms of that given speech

community which are also reflected in the RO sets. It is also very important to bear in mind that markedness is a gradient concept, and not a categorical one (Myers-Scotton 1989: 335), meaning that it is not a binary distinction between marked and unmarked choice.

The markedness theory follows in the Gricean tradition of the cooperative principle and conversational maxims (Grice 1975), and holds that for each linguistic exchange there exists a marked and an unmarked choice of expression in relation to the speech community's linguistic norms. The unmarked choice refers to the normative set of rights and obligations, i.e. the standard and community-wide accepted linguistic behavior, while the marked choice refers to the substandard set of RO. Each communicative situation involves the "negotiation of identities" (Myers-Scotton 1983: 115) and relationships, which means that through the choice of linguistic expressions and codes, interlocutors pit their identities against each other and through the observation or non-observation of the maxims generate conversational implicatures regarding the rights and obligations. A marked choice calls for the readjustment and recalibration of the expected social distance between the participants of the conversation. Code choice also reflects and depends on the interpersonal relationship of the participants of the conversation, and each exchange is associated with a marked and an unmarked set of rights and obligations.

Where people use a mixed language or engage in code-mixing regularly in a multilingual environment, CS is the norm (i.e. it is the unmarked choice) and the status quo. Whenever code-mixing is not normative (i.e. it represents the marked choice), it involves the negotiation of rights and obligations between the speaker and the addressee. It is also possible that the external circumstances of the exchange necessitate switching but the speaker refuses to do so, thereby opting for the marked choice of not switching instead of the unmarked choice of switching, thereby indicating the power distance between themselves and the addressee.

Myers-Scotton in general seems to believe that CS is a conscious and deliberate choice on the speaker's part, and that speakers are always aware of the norms and maxims of the speech community. This is reflected by her treatment of code-switching as a process whereby speakers are directed by the negotiation principle to make choices, and also by the previously cited and discussed view of code-switching being speaker-motivated (Myers-Scotton's 1989: 336). To this end, she lays down four types of motivations for engaging in code-switching (ibid. 336-339). One motivation can be, as it has been discussed in the previous section, speaker-external, namely that the context in which the exchange takes place has altered and participants must accommodate for this change in their linguistic behavior, and not switching

would constitute opting for the marked choice. Another type of motivation is when switching represents the unmarked choice in the exchange, and no speaker-external forces necessitate the switch. In such cases CS represents the norm, typically among bilingual interlocutors in informal, ingroup contexts. Thirdly, code-switching can represent the marked choice in certain contexts, mostly as a way of signaling a change in social distance, often as a cohesive force of in-group solidarity or to increase the power distance between the interlocutors. Finally, switching can also be an “exploratory choice” (Myers-Scotton 1983: 338), whereby the speaker explores the situational and social identity factors of their interlocutor upon encountering each other for the first time. While these motivations for engaging in CS are compelling and logically sound, they imply that code choice is deliberate on behalf of the speaker, which is in itself mostly true, but it should also be factored in that it is often dictated or at least influenced by social factors.

The markedness theory forms part of perhaps the most influential theory of code-switching that Myers-Scotton (1993a) put forward, which is the Matrix Language-Frame model (MLF). In this dissertation the MLF features less prominently than the markedness theory. The MLF is a model for intra-clausal code-switching, and proposes the existence of a Matrix Language (ML) and an Embedded Language (EL) in the code-switching speaker’s mind, and builds on the observation that the two languages that participate in CS are not equally represented in mixed utterances and clauses. In essence, the Matrix Language is the dominant language which supplies the morphosyntactic frame of constituents (Myers-Scotton 1993a: 82), the main morphological and grammatical structure, into which elements from the Embedded Language can be inserted.

Instead of establishing structural constraints, at the core of the Matrix Language-Frame model are two principles based on this imbalance of languages. These are the morpheme order principle and the system morpheme principle, which apply to mixed constituents. The morpheme order principle states that in ML + EL constituents containing at least one EL element, the surface order of morphemes will be corresponding to the morpheme order of the ML (Myers-Scotton 1993a: 83), and the system morpheme principle means that in ML + EL constituents all system morphemes will come from the ML (*ibid.*). In her model, Myers-Scotton makes a differentiation between content morphemes and system morphemes, where the latter corresponds to the bound (inflectional, derivational) morphemes and function words of a language. The system morpheme principle would entail that no EL function morphemes could be inserted into mixed constituents (cf. Muysken 2000: 16).

In relation to the topic at hand in this dissertation, the MLF model can be utilized to investigate, especially with the help of the system morpheme principle, which of the two languages in a hybrid toponym likely functioned as the Matrix Language, which could reflect the dominant language in a given area. However, in congruent lexicalization the overt separation of languages is not necessary and not feasible. A possible caveat of utilizing the MLF for the analysis of hybrid toponyms is that there are only a maximum of three elements in each place-name, and Old English shared its naming practices with Old Norse, therefore the morpheme order principle will not be able to determine which is the Matrix Language. This is true in the case of English-Scandinavian hybrids because English and Norse place-names are structured in the exact same way (generic + specific), but the principle can be useful in the case of Celtic names, because Celtic place-names show an order of elements that is reversed as compared to the English order⁴. In the case of those hybrid names which have either two Norse elements or a Norse and Celtic element and the order of elements follows the Celtic model (the so-called Hiberno-Norse inversion compounds⁵) this principle would entail that Celtic was the matrix language.

The Matrix Language-Frame model and the markedness theory have the following implications concerning hybrid place-names in general and the ones in the focus of the present dissertation. The markedness theory is most probably manifested in hybrid place-names, especially the Anglo-Norse ones, as an unmarked choice of the second type as outlined above. For this to be the case, we need to adopt the narrow definition of hybrid toponyms from Section 2.1, namely that those should be considered hybrid ones which emerged in a bilingual context and not through borrowing. In the case of the English-Scandinavian hybrid names, the motivation was not external in the sense that no external force necessitated the continuation of code-switching behavior because OE-ON CS was not present before. It is not of the third type either, because it could not signal in-group solidarity given that one of the groups of speakers in the contact situation conquered and ruled over the other. Finally, it could not have been of an exploratory nature either because that would again presuppose the existence of a history of OE-ON code-switching. The situation then is likely to have been that CS represented the unmarked choice in a bilingual community.

⁴ For instance *Aberdeen* ‘mouth of the river Don’ in which *aber* is the generic and means ‘mouth of a river’ (Mills 2011: 2). In the English order of elements, the name would be along the lines of **Deenaber*.

⁵ For instance *Kirkoswald* (CUMB) ‘church dedicated to St. Oswald (7th century king of Northumbria)’ < ON *kirkja* ‘church’ + OE personal name (Mills 2011: 280), with a Celtic order of elements. The English and Norse order would yield **Oswaldskirk*. A more thorough analysis of such names and their implications for the Matrix Language-Frame model will be presented in Section 2.2.1.

In the case of Celtic-English hybrid names the plausible scenario was one involving borrowing from the get-go as the Germanic settlers “received” an already existing set of toponyms, which they adopted and integrated. However, this is also true for the initial phase of the Norse settlement, but many of the English-Scandinavian hybrid names are established on hitherto uninhabited areas, especially names in *-þorp*, which indicate that they were established as new settlements at a later period. However, in many cases in the MLF model it is difficult to determine which of the two languages functions as the ML, for which various diagnostic devices have been proposed (cf. Muysken 2000: 64-67 with an overview of a number of criteria). Determining the ML in the case of hybrid place-names poses extra problems, because instead of a whole stretch of language we only have a two or three-element name, and all we can rely on is the dominant language of the speech community being the matrix language.

Much of the Matrix Language-Frame model’s implications for hybrid toponyms have already been discussed above. Still, in conclusion it can be said that the MLF model is relevant for hybrid names because in the case of Celtic-English and French-English hybrid names English can be established as the Matrix Language, with Norse-English hybrids the cases is less clear. The great benefit of the MLF model is that it can accommodate switching occurring in constituents smaller than the sentence, therefore hybrid place-names (which are in essence nominal or attributive compounds) can also be considered to be instances of code-switching. However, the Matrix Language-Frame model, while in theory would permit the analysis of hybrid place-names, was created for the analysis of sentences, therefore its implications have to be adjusted for the examination of hybrid toponyms. Ideally, hybrid place-names would require the development of a separate framework of analysis (cf. Section 2.1 about the scarcity of their treatment in literature) which would require extensive research and the establishment of verifiable and falsifiable theories that lies far outside the scope of this empirical dissertation.

2.2.2. Muysken’s (2000) typology of code-mixing

In his treatment of bilingual speech and code-mixing, Muysken (2000) formulates an approach to the study of CM from an angle that is different from Myers-Scotton’s primarily sociolinguistic and pragmatic model (i.e. the markedness theory), but bears resemblance to her structure-based Matrix Language-Frame model, in that Muysken’s approach also focuses on the grammatical aspects of language contact and the patterns of code-mixing.

Muysken (2000: 3) introduces a three-way classification and typology of intra-sentential code-mixing, whereby he establishes the categories of insertion, alternation and congruent lexicalization. Insertion refers to the type of code-mixing whereby material from one language is embedded (or inserted) into the structure of another language. In connection with insertion, Muysken (2000: 60) hypothesizes that “borrowing, nonce-borrowing⁶, and constituent insertion all fall within the same general class and are subject to the same conditions” therefore they can all be described by the category of insertion. In the case of insertional code-mixing a constituent, which could even be a single constituent, is inserted into a frame of another language. In this sense “constituent” is defined as a syntactic unit that is either a lexical item or a phrase. Muysken (2000: 63-64) identifies five main characteristic features of insertion: (i) the vast majority of insertions are single constituents, (ii) insertions appear in a nested “A B A” structure, meaning that the inserted elements from language B interrupt a grammatical sequence of constituents from language A and the constituent from language A enveloping the insertion are grammatically related to each other, (iii) the switched elements are more likely to be content words than function words, (iv) in the case of switching syntactic constituents the switched elements are objects or complements instead of adjuncts, and (v) insertions are morphologically integrated, complete with case markings.

Alternation means that the involved languages remain separated and structures coming from the languages occur in an interleaved, alternating pattern with each other. The following four features of alternation can be identified: (i) a sequence of several constituents are switched, (ii) the switch occurs in a non-nested “A ... B ... A” sequence where A and B refer to languages, instead of constituents as in the case of insertion discussed above, meaning that we are dealing with longer sequences and stretches of language here, (iii) as a direct consequence of the second criterion, alternation is characterized by the use of multi-word, complex structures, and (iv) functional elements, discourse particles, and adverbs can be incorporated into a language via alternation (Muysken 2000: 96-98) whereas insertion typically involves nouns, adjectives, and other content words.

Finally, congruent lexicalization occurs when two languages share the basic grammatical structure of a sentence, into which elements of vocabulary from either language can be inserted (Muysken 2000: 127). Congruent lexicalization (CL) can also be defined as the lexicalization of “material from different lexical inventories into a shared grammatical

⁶ The term “nonce borrowing” was introduced by Poplack (Poplack, Sankoff & Miller 1988) and it is taken to refer to those borrowings that occur only once and have not become integrated into the recipient language, as opposed to widespread loans. Such items are classified as borrowings instead of instances of CS because otherwise they would violate the constraints of code-switching.

structure” (Muysken 2000: 3). This type of CM happens when closely related and highly similar languages of roughly equal prestige are affected and there is no overt language separation (Lipski 2009: 2-3). Congruent lexicalization can also take place in a setting where a dialect mixes with the standard, which could also be considered a form of style shifting, and is therefore in accordance with Myers-Scotton’s (1989: 336) view of code-switching, outlined in Section 2.1, that permits the mixing of different dialects and varieties of the same language. This is also echoed by Muysken (2000: 122) who proposes that “code-mixing is not different, in principle, from variation.”

Congruent lexicalization often emerges in a bilingual setting, and can be triggered by the following two factors. Firstly, CL is brought about by an “overabundance” of homophonous diamorphs⁷ in the two languages or varieties, and secondly, CL is triggered by structural equivalence, but lexical correspondence is not a prerequisite of congruent lexicalization, and it is also possible that only one of these triggers are present to actuate CL. In congruent lexicalization, both languages “contribute to the grammatical structure of the sentence” and the vocabulary can come from different languages or it might be shared as well (Muysken 2000: 127). This means that the grammatical structure shows significant overlaps and the vocabulary can show overlaps between the two languages.

The following features of CL can be identified (Muysken 2000: 128-135): (i) there is linear and structural equivalence between the two languages due to the fact that they are syntactically identical⁸, (ii) similarly to alternation, but due to different reasons, CL allows the occurrence of multi-constituent mixing, (iii) non-constituent mixing can also occur in congruent lexicalization, which means that not full constituents are mixed, (iv) the mixed structures in CL occur in a non-nested “A B A” pattern, meaning that the constituents from language A enveloping the foreign element are grammatically not related to each other, as opposed to the nested configuration in insertion (cf. also Lipski 2009: 26), (v) all lexical categories, even function words can be switched, (vi) bidirectional mixing is also possible in congruent lexicalization, referring to the possibility of inserting elements from language A into language B and vice versa, something which occurs quite infrequently in insertion and

⁷ Based on Crystal’s (2011: 141-142) definitions of a diasystem, that is the “network of formal relationships which shows the common linguistic system assumed to underlie two or more dialects,” and a diaphone, i.e. “an abstract phonological unit set up to identify an equivalence between the sound system of different dialects,” a diamorph can be defined as an abstract morphological unit that displays equivalence between the morphological system of different dialects (or languages).

⁸ I believe this criterion to be true at least in the case of different varieties undergoing CM, but the syntactic structure of more distant yet still closely related languages may not necessarily be identical even though they would still participate in congruent lexicalization.

alternation⁹, and (vii) morphological integration and word-internal switching are also commonly occurring phenomena in CL due to the similarity of the involved systems.

As it was discussed previously, lexical borrowing is most frequently associated with insertion, however congruent lexicalization can also provide a frame for loanwords. Muysken (2000: 150-151) briefly describes English-German hybrid (or in his terminology mixed) compounds which he considers to be word-internal borrowings that came about through congruent lexicalization and ascribes to the effect of the two languages having a shared word grammar (i.e. morphological structure and word-formation rules). Muysken suggests that in such mixed compounds, the matrix language is the one which supplies the head of the compound word. In the previous subsection on the Matrix Language-Frame model it was mentioned that in the case of hybrid place-names it could be difficult to identify the ML with certainty, and relying on identifying the language of the head might prove to be useful. In the case of toponyms, which in this dissertation are treated as compound words (see Chapter III), generics are considered to be heads.

Table 1 below summarizes the main features and diagnostic devices of Muysken's categories of code-mixing. The shaded cells represent those features that are relevant for the analysis of hybrid place-names.

insertion	alternation	congruent lexicalization
single constituents (objects, complements, or single words)	several constituents	multi-constituent & non-constituent mixing
nested structure	non-nested, interleaved structure	non-nested mixing
content words (nouns, adjectives)	functional elements, discourse particles	any lexical item
morphological integration	unintegrated long, complex fragments	morphological integration & word-internal CM (presence of homophonous diamorphs)
unidirectional	dominantly unidirectional	bidirectional

Table 1. Summary of Muysken's (2000) code-mixing typology

⁹ Muysken provides the figure of typically nearly 100% of insertions and 75% of alternations being unidirectional in code-mixing corpora, and in the overwhelming majority of cases, elements from the dialect are inserted into the standard variety and not the other way round. In Muysken's framework of congruent lexicalization, (v) and (vi) are also considered to be made possible by the lack of a matrix language in congruent lexicalization that would determine the language of the function words (hence feature (v)) and the frame for the embedded language (hence feature (vi)).

It can be argued that the situation presented here in the case of English-Scandinavian place-names is in fact a manifestation of congruent lexicalization, which is not of an intrasentential nature, but word internal or intralexemic instead (cf. Muysken 2000: 137). The close relatedness which is a requirement of the operation of CL, and the roughly equal prestige, which contributed to the emergence of such a high number of hybrid place-names, was very much the configuration in the case of the contact between English and Norse. At the time of their contact during the period between the 8th and 11th centuries, these two languages were very much similar¹⁰, with a degree of mutual intelligibility, and as will be discussed in Chapter IV and V they were in an adstratal relationship, even if partially. The fact that the OE and ON were close relatives of each other with structural, morphological, and lexical similarities entails that homophonous diamorphs could frequently occur in these two languages. Furthermore, the property of bidirectionality is also relevant for OE-ON hybrids, because English specifics occur with Norse generics in roughly equal numbers as Norse specifics occur with English generics, which can be taken to represent the bidirectional nature of hybridization and mixing. Concluding the discussion of congruent lexicalization, it can be mentioned that perhaps the most important aspect of CL for hybrid place-names is that it is grammatically unconstrained, and that “linguistic convergence feeds into congruent lexicalization” (Muysken 2000: 221) which is especially salient in the case of English-Scandinavian hybrid toponyms.

Finally, Celtic and French hybrid names likely conform to the category of insertion because they are the ones which most likely came about through the borrowing of Celtic and French toponymic elements, and because the phenomenon of insertional code-mixing bears significant resemblance to lexical borrowing (Muysken 2000: 69), especially nonce-borrowings. The unidirectionality of insertion is also reflected in the place-names data of the corpus, as there are far more Celtic-English hybrids (N = 196) than English-Celtic ones (N = 2). Finally, the fact that single constituents and even single content words can be utilized in insertion points in the direction of insertion (and CL for English-Scandinavian hybrids) and not alternation which makes use of functional elements. However this analysis might be inconclusive given that the investigation is not focusing on entire sentences but multi-element

¹⁰ There were, however key phonological differences which are reflected in some of the Scandinavianized forms, such as the substitution of [sk] for [ʃ] as in Fiskerton (‘fisherman’s village’ containing OE *fiscere* ‘fisherman’ with [sk]), or [k] for [tʃ] as in Digby (< ON **diki-bý* ‘village at the ditch’ with [k]), to name two of the most frequently occurring ones, among others. Due to the lack of space here, for a detailed, comprehensive treatment of phonological differences see Townend (2002: 31-41, 61-63).

words. Also, English-Celtic (e.g. *Yorkshire* and *Devonshire* cited previously) and English-Latin hybrid names are extremely low in occurrence (especially compared to English-Norse hybrids) and only a few recurrent generics are used in them, albeit with high frequencies, so it can be concluded that they most likely arose through the borrowing of said generics which were then used by speakers of English to create the names. After all the data of the corpus has been surveyed, it will be explained in Chapter V and VI whether the situations at hand indeed conform to the category of insertion and congruent lexicalization or not.

2.3. Historical code-switching vs. lexical borrowing

In their essence, code-switching and borrowing both function as indicators of social and linguistic contact between groups of speakers, although this contact is not necessarily of a physical nature. In a historical perspective it is very difficult to establish a demarcation line between the two phenomena, mostly because CS is often the precursor of borrowing, and also because the incidence of CS is higher in spoken language and apart from the past century, all the available historical sources in which CS occurs are exclusively written ones. This means that our present-day knowledge of historical code-switching is based only on written documents which is somewhat problematic, because written language is characteristically more rigid and conservative than the spoken variety. This would hold true even in an era before standardization and the emergence of the idea of normative linguistic behavior. As a result, our knowledge of the frequency and incidence of historical CS and CM is skewed. A further factor complicating the picture in the case of place-names is that the majority of the English toponyms are first recorded only in the Domesday Book of 1086, and only a comparably small proportion of them can be found in other, earlier sources, which are mostly charters. Related to the issue of sources for toponyms and other linguistic material, another serious problem presented by historical sources of English is the scarcity of surviving manuscripts from many dialects that could shed more light on the early forms of English and also on place-names. In this section the possible distinctions and similarities between (historical) code-switching and lexical borrowing will be surveyed briefly.

In one of the earliest treatments of borrowing, Haugen (1950: 211) also talks about language mixing and switching “rapidly from one [language] to the other” which could be understood as a description of code-switching. In the same paper, Haugen distinguishes three main types of borrowed elements, depending on whether they incorporate the foreign material or substitute it with native terms: loanwords, loanblends, and loanshifts. Loanwords operate

with “morphemic importation without substitution,” meaning that the form of the foreign word is completely taken over. The category of loanblends covers those elements which show morphemic importation as well as substitution, which means that some part of the words are taken over as they are, and some are substituted for with native elements. Therefore, hybrid compounds and other hybrid words belong here. Finally, loanshifts can be equated with loan translations or calques, and they exhibit morphemic substitution without importation, i.e. in these cases only the meaning of the original word is taken over and not the form, as that is expressed via native surface forms. The importance of this categorization from our perspective is that as early as Haugen’s analysis, hybrid linguistic items were recognized as a possible outcome of bilingualism (in Haugen’s paper) which was later broadened to language contact and linguistic interference (e.g. Thomason 2001: 153).

A universal developmental path of loanwords as opposed to code-switches is that in the recipient language they gradually assimilate into its phonological, morphological, and syntactic subsystems and except for their etymological background and occasionally some phonetic characteristics they do not retain any of their original features (Poplack 2015: 919). The original meaning of loanwords is also taken over, although elements in the recipient language have fewer meanings than in the donor language, because they are used less frequently there (Mańczak 1985: 374-375). Code-switches on the other hand retain their original grammatical properties, and their meaning does not undergo narrowing. As opposed to borrowing and other contact-induced changes (e.g. pidginization, creolization), code-switching does not involve the replacement and transformation of the existing vocabulary and grammatical features of a language either (cf. Sankoff & Poplack 1981: 4). This true at least in the short run, as frequent code-switches in the long run can become established as loanwords and may even lead to the replacement of native lexical items.

Myers-Scotton (1989: 340) discusses the three main ways in which synchronic CS differs from borrowing. Firstly, starting off with a similarity, they are both “normatively sanctioned”, meaning that from the perspective of the norm, their occurrence is allowed in language, the entire speech community recognizes their use, albeit not always positively, especially purists, who are very much against CS and the usage of foreign words in any form¹¹. Poplack (1980: 585) reflects on this too, describing code-switching as an “overtly

¹¹ The purists’ battle against the acceptance of loanwords into the language is very much synchronically grounded, and ultimately futile. It is in the synchronic slice of the language that foreign words and elements stand out the most, because at that point they have not yet been integrated (and many foreign elements feature in the language as code-switches and are unlikely to gain any currency), but with the passage of time they all become phonologically, morphologically, and syntactically fully integrated into the receiving language.

stigmatized sociolinguistic marker.” The difference between CS and borrowing, however, lies in the fact that in borrowing *specific words* are taken over, which then become part of the recipient language’s lexicon and are normativized, while in code-switching instead of discrete lexical items it is the *behavioral pattern* of using words of the embedded language that is normativized, at least in those speech communities where CS is widespread.

Secondly, the frequency of borrowed elements and exactly what elements can occur is predictable based on the speaker’s knowledge of the matrix language, but what elements from the embedded language will occur as code-switches is unpredictable. Borrowed words have a relatively stable status in the current stage of language, while code-switches are very often ephemeral. Naturally, this does not mean that once a loanword enters a language it will never fall out of use, as it is evidenced by Scandinavian loanwords which were ousted by Norman French terms after the conquest of 1066 (Lutz 2013: 567). Also, perhaps most importantly from our perspective, borrowed words are available to monolingual speakers of the recipient language, while code-switches are not as their use presupposes at least some bilingual competence (cf. Pfaff 1979: 295-296, Muysken 2011: 304). This feature of CS is often held to be the clearest yardstick for differentiating it from borrowing.

Thirdly, only CS can “convey micro-level social meaning in interpersonal negotiations,” meaning that only code-switching is capable of indicating in an exchange the prevalent rights and obligations, and only CS can invoke a change in the rights and obligations¹². Borrowed words do not bring about a negotiation of identity, and do not bring about implicature, although, recent loanwords in a language’s lexicon might trigger such a response from addressees.

Muysken (2011: 302-303) also sets up three criteria by which code-switching could be distinguished from borrowing, which I have already touched upon in the discussions of CS in various sections of this dissertation. These are (i) the adaptation of the foreign element to the matrix language, (ii) the bilingualism of the speech community, and (iii) the amount of material taken over from the donor language. Muysken (ibid.) also poses the question whether it is useful at all to try and distinguish between these two phenomena. Muysken also formulated this concern earlier (cf. Appel & Muysken 2005: 173) and argued that “it is not possible to distinguish individual cases of code-mixing from not-yet-integrated borrowings on the basis of simple diagnostic criteria.” Appel & Muysken (ibid.) cite the distinction between

¹² For an explanation of these terms see Section 2.2.1.

the *use* of two systems, simultaneously as in CS and CM, and the *adoption* into a system, on an item-by-item basis as in borrowing, as the main underlying theoretical problem here.

In the case of borrowings, the borrowed elements very often fill lexical gaps in the recipient language and are taken over because the language lacks an equivalent term. Borrowing can also be initiated if an equivalent does exist but it might not be used by the entire community or it might not be known to the speaker (cf. Pfaff 1979: 297). A traditionally made distinction in the realm of loanwords is between necessity borrowings and luxury borrowings. Necessity borrowings are the ones that were described above, i.e. words taken over in lack of an L1 equivalent, mostly in order to denote new or previously unknown concepts or things. On the other hand luxury borrowings are the results of the donor language having a prestige that is perceived to be higher than that of the recipient language. They always have an already existing and established equivalent in the borrowing language, which means that they are economically superfluous, hence their name. Luxury borrowings do not serve the purpose of filling a lexical gap, they enrich synonymy and add stylistic value. However, they can coexist with genetically inherited lexical items with the same meaning (and may lead to the formation of etymological doublets if both the native and the borrowed element are the reflexes of the same etymon), they may restrict the meaning and use of native terms or entirely drive them out of use. Necessity borrowings can also reflect the perceived superiority of the donor language and/or its speech community because instead of coining a term utilizing the word stock and word formational rules of the recipient language, speakers opt for taking over the foreign item because of its higher prestige.

We can talk about necessity borrowings in the case of borrowing already existing place-names into the language of a newly settled speech community, as the newcomers do not yet have a toponym for the existing settlements, therefore it is logical that they use the ones given by the population that settled before their arrival. This is how Celtic toponyms possibly entered the English language.

Code-switching can be related in its purpose both to necessity borrowings and to luxury borrowings. Switched items can be included in a stretch of language in order to express a concept or thing that the other language has no term for or a circumlocution would be cumbersome (as in necessity borrowing). On the other hand if the language does already have a word for a given concept, then CS can be used to allow the speaker to express their identity or use a term that they can retrieve from their mental lexicon more easily or attribute higher prestige to (as in luxury borrowing). This is also the case in Muysken's (2011: 303) discussion, as in his view the question to ask in determining the status of a foreign element in

a language is whether the inserted item denotes a “new concept”, in which case we are likely to be dealing with a borrowing, or does it “duplicate and already existing word” in which case code-switching is the more plausible explanation. This approach has implications for hybrid place-names as well, namely that many of the Scandinavian generics found in hybrids lack an English equivalent, which would mean that based on this criterion alone these generics were borrowed. However, due to their low frequency of occurrence they could also be classed as code-switches. In the case specifics on the other hand we can find numerous overlaps and similarities between OE and ON items both in terms of form and meaning. Also, the Scandinavian generics occurring in hybrid place-names do not describe a new concept, as loanwords would, but add a new word for already existing concepts (e.g. ‘farmstead’, ‘clearing’, ‘wood’, etc.) as luxury borrowings or code-switches would. This problem illustrates that code-switching and borrowing are often inseparable and that their distinction based solely on binary features is ill-conceived (cf. Pfaff 1979 who also considers the boundaries between code-switching and borrowing to be vague and fuzzy). Instead, the exact nature of the situation and context in which CS and borrowing takes place should be considered when assessing the outcome of the interference.

In his unified approach, Backus (2015) calls attention to the fact that the distinction between CS and borrowing is a very much synchronically grounded one that researchers have tried to pin down in the light of the constraints discussed in the previous section. This state of affairs is untenable, he maintains, as the constraints are often inherently flawed and circular, and do not take diachronic aspects into consideration. The key notion of Backus’ theory in relation to borrowings is connected to the frequency of occurrence that I have also advocated and put forward previously in this dissertation.

The starting point is that linguistic units have different degrees of entrenchment in each individual’s mind and mental lexicon, meaning that they are imprinted with different strengths, and repetition increases the depth of entrenchment. Full sentences and complete utterances are very unique, and very unlikely to occur more than once, due to the open-endedness of language, and are thus unlikely to be entrenched at all. Words, combinations of words, and various word-forms on the other hand show “wildly fluctuating” frequencies (Backus 2015: 22) and thus equally fluctuating degrees of entrenchment. Code-switching is the synchronic phenomenon that has borrowing as its diachronic result. Synchronically, code-switches, as they occur unpredictably (cf. Myers-Scotton (1989: 340) as discussed above) and are very often used on a one-off basis, are very slightly or not at all entrenched. For a word to enter the lexicon of a language from the individual speaker’s linguistic repertoire and idiolect,

extensive usage and thereby deeper entrenchment is required. Tentatively, code-switching and borrowing can be imagined to be different stations on a continuum, with insertional code-switching representing one of the end-points (in synchrony) and lexical borrowing the other (in diachrony). Thomason (2001: 136) also reaches a similar conclusion, namely that in essence, code-switches are foreign elements and structural features that are introduced into a language, and once they have entered the recipient language they progress along the same developmental path as language-internal changes and innovations would. This progress and spread is achieved via repeated use and exposure to more and more speakers, hence the deeper entrenchment.

Before concluding this chapter, a few words shall be in order about the relationship of borrowing and stratal influences. Vennemann (2011: 220) describes the typical scenarios conditioned by the stratal relationship of the affected languages in which borrowing can take place. From substrate languages only a small amount of loanwords are taken over with the exception of toponyms, as the main motivation in this case is to have words for the cultural and environmental aspects of the other group. In the case of superstrate borrowing, however, the motivation is to communicate with the elite, and the perceived prestige of the elite and their language can bring about quite extensive lexical borrowing. Vennemann (*ibid.*) introduces a third type of stratal influence, namely prestratal by which he refers to cultural borrowings like the Greek and Latin loanwords in modern languages. According to Vennemann's definition, a prestrate is a language that "influences another language on account of its cultural prestige or appeal" for which physical contact and sharing a territory is not a prerequisite. Although Vennemann does not cover adstratal influence, he does provide a characterization of an adstrate, which is somewhat problematic and circular, as it is defined simply as a language that is neither a substrate or a superstrate. A better candidate for the definition of adstrate can be that of Tristram (2007: 198), who also cites the relationship between Old Norse and English in the Danelaw as an adstrate, according to which "[a]dstrates exert influence on their adstrates on all levels but mostly on their lexicon"

There are two key concepts introduced by Vennemann (2011: 222-224) in connection with linguistic interference and language shift that are relevant for the dissertation: Lexical importation and structural importation. In general, importation is considered to be the residue of the second language learning process that is undertaken by shifting speakers, and the resulting forms and features in the target language are those forms and features of the source language that survived the language shift (*ibid.*). The terms are quite straightforward, lexical importation means that source language expressions or their meanings (cf. the category of

“loanshifts” in Haugen’s (1950: 211) terminology) are continued to be used in the target language despite the fact that shifting speakers are trying to speak the target language. Structural importation on the other hand means that structural features of the source language are carried over by substrate speakers. When speakers of the substrate shift to the superstrate (substratal language shifting) lexical importation is not so pronounced but it can be accompanied by significant structural importation, and when superstratal language shift occurs, it is preceded by heavy borrowing from the superstrate into the substrate.

Finally, as a way of recapitulating this section’s main message, Thomason (2001: 133-136) outlines the following three ways in which code-switching and other linguistic interference phenomena can be manifested that are of interest to us. First, if a foreign element gains currency in a monolingual speaker’s language use, then we can assume that it has become an “interference feature,” in other words, a borrowed item. Secondly, code-switched items are not integrated into the recipient language, whereas borrowings are, in Thomason’s terminology, nativized. Third, if a foreign element occurs only once in a bilingual’s discourse, then it is likely to be an instance of CS rather than borrowing. With all this in mind, Thomason raises awareness to a number of pitfalls and caveats concerning making a distinction between CS and borrowing. For instance, in a context where every speaker is bilingual the first criterion falls flat, or certain words that are clearly borrowed, such as personal names, might be pronounced according to the donor language’s rules and thus not be nativized, or the fact that measuring frequencies with absolute certainty is a hopeless task.

Concerning linguistic interference in general, Thomason & Kaufman (1991: 35) point out that the main determinant of the outcome of language contact is the speakers’ sociolinguistic history and not the structure of their language. Linguistic and structural constraints all fail in favor of social and cultural factors conditioning the outcome of language contact. They argue that the social factors determine the direction of the interference, i.e. which language borrows and which supplies the forms borrowed, also the extent of the interference, i.e. which layers are affected and how deeply, and finally what kind of features are transferred between the languages. The explanation may seem banal, but language change, and contact induced language change, as Thomason (2000) rightly observes, is largely unpredictable.

In conclusion, code-switching is quite an elusive concept to grasp definitively, as it has been analyzed from a variety of perspectives, and it lacks a clear definition (cf. also

Gardner-Chloros 2009: 10-13¹³). Borrowing and code-switching are strongly related to each other, and are perhaps best conceived of as points on a continuum and not in absolute terms (e.g. Pfaff 1979, Muysken 2000, 2011, Backus 2015). As it has been demonstrated before, code-switching can occur within words in the form of hybrid compounds (Muysken 2000: 151), and between stems and affixes of complex words (Thomason 2001: 135), which bears significant relevance for the analysis of hybrid place-names as instances of word-internal code-switching.

¹³ For a synthesizing and broad overview of what code-switching and CS research can encompass see Gardner-Chloros 2009: 1-19, and for a briefer treatment see Bullock & Toribio 2009a).

III. The Genesis, Morphology, and Semantics of Place-Names

After having defined and surveyed the sociolinguistic framework of the analysis in the previous chapter, in this chapter the linguistic background of place-names and their formation will be brought under scrutiny. The standpoints presented in this part hinge on the underlying assumptions that (i) place-names originate in language as semantically and formally transparent lexical items, (ii) toponyms can be construed as darkened compounds, and (iii) due to the so-called onomastic divorce that affects the semantic characteristics of proper names certain linguistic changes might affect them differently and/or to a greater extent than free lexical items.

As a result of treating toponyms as if they were compounds, it is also assumed that they share certain developmental characteristics and semantic properties with regular compound words. There are two very important features that are to some extent shared by these two types of words which should be mentioned here as preliminaries. One is the fact that the reference of compounds is more singular than that of free words, and place-names (and other proper nouns) are entirely monoreferential, that is they refer uniquely. The other feature, related to this restricted referentiality, is that free words in compounds undergo semantic isolation and place-names undergo onomastic divorce, both of which affect how they relate to free lexical elements. These properties, as it will be argued below, can make compounds and toponyms more susceptible to obscurity and other changes due to their divergence from the main lexicon and the main lexical meanings of the free counterparts of their constituents.

Concerning the structural, and especially morphological, makeup of compounds and toponyms, their similarities will be highlighted, drawing attention to the various historical changes that affect them phonologically, morphologically, and orthographically. This analysis ties in with and is built on the previously introduced semantic framework dealing with the obscurity of meanings. How the morphological-structural processes correlate with the semantic ones will also be discussed. Due to the fact that folk etymology plays a role in the reanalysis of obscured words, it will also be touched upon briefly in this chapter.

3.1. The appellative origin of place-names

Every name-giving act is a semantically conscious and deliberate one, and at least to some extent every name is semantically motivated¹⁴ (Ditrói 2017: 24-25). Therefore, place-name formation cannot be considered a random, haphazard phenomenon, because all proper nouns ultimately originate as semantically transparent, meaningful lexical items, adequately descriptive of whatever they denote, therefore the same phonological, morphological, word formational and syntactic rules and constraints apply to them as to the whole of the language and to the formation of words belonging to other word classes (see Nicolaisen 1976: 151, Clark 1992a: 452, Gammeltoft 2007: 481). These names then continue to be transparent on the formal and on the semantic level as well for long periods of time, but are then very often obscured and rendered opaque by folk etymology and/or the regular, historical changes that take place in the language¹⁵.

Nicolaisen (1974) distinguishes three semantic levels of naming which are (i) the dictionary meaning of the name's constituents (lexical level), (ii) the reason or motivation for coining the given name (associative level), and (iii) the meaning of the name as a name (onomastic level). In the genesis of a place-name each of these three tiers is represented and during the development of any toponym the first two might undergo certain changes that result in the semantic obscuration of the given name. The common nouns that make up those place-names which are not derived from personal names can gradually lose their appellative meaning in the given place-name and begin to refer to only one unique entity, i.e. the locality, a process which Blanár (2009: 89) termed "onymisation". However, it cannot be determined at exactly which point during its development the appellative origin of a place-name or a personal name becomes obscured. Furthermore, this process does not mean that all names from the moment they are created are by definition always completely devoid of any meaning and their sole purpose is to function as mere labels used for the identification of localities. Names do carry meaning during the phase when they are still transparent and analyzable (cf. Clark 1991: 284, Ditrói 2017: 13).

¹⁴ cf. also Colman (1991: 13) who asserts that "Old English personal names are formed from elements cognate with common words." In Colman's understanding of names, name-elements are considered to be *cognates* with free lexical items rather than to be free lexical items that are used in names because name-elements behave and develop differently than common words.

¹⁵ The example of the development of the name *York* is a well-known and frequently cited one, and it provides quite a good glimpse into the various processes that affect toponyms and the ways in which various layers of settlers and inhabitants treat the place-names they encounter (for an in-depth analysis see Fellows-Jensen 1987).

Onymisation should not be a bar for the appropriation and adaptation of foreign settlement names either, because the etymological background and, more importantly, being able to attribute lexical meanings to semantically opaque place-names are irrelevant factors for speakers (as argued by Dalberg 2008 [1977]: 52), and by extension, being familiar with the etymological background of any word that is not a name is also irrelevant for the average language user. Nicolaisen (1974: 104-105) also makes a similar point, namely that the lexical meaning of any name is an irrelevant factor for the speaker, even if the name in question is still transparent and the lexical meaning is accessible. Folk etymology and its related processes of reanalysis and analogy are especially salient factors in making sense of the obscured toponyms and lexical items. In this dissertation, Coates' (1987: 321) definition of folk etymology is accepted, namely that it is an analogical reformation and a "formal change whereby one element of a morphologically complex form comes to resemble more closely, or become identical to, some other morph in the speaker's lexicon." This process during which the morphological forms of two elements converge to each other is reflected in place-names by the "replacement of alien elements by similar-sounding and more or less apt familiar ones" (Clark 1992a: 479-480).

Names (and to some extent common nouns and other 'non-proper' types of words) do not have to be transparent and readily analyzable because if speakers cannot make sense of certain names (or nouns, etc.) they are very likely to resort to the practice of folk etymology which helps them connect the unanalyzable item with a familiar, analyzable one which often results in an erroneous interpretation, one that is perceived to make sense or at least enables the speakers to attribute a meaning to the word in question. The occurrence of place-names in everyday speech is likely to be quite infrequent but nevertheless they do not enjoy any special position, treatment or attention in the average speaker's language use. Lastly, place-names and personal names are usually very strongly tied to personal and group identities, therefore attributing certain perceived qualities to them is quite a commonly employed practice (cf. Nicolaisen 1978 about connotative names).

A usual distinction or differentiation is made between the lexicon and the onomasticon, and a full treatment of this issue would be justifiable if the questions at hand were looked at from more of a theoretical perspective. I maintain, however, that in the context of the present dissertation it is rather unnecessary to deal with aspects of their distinction in greater detail, as my investigation is from a more practical and usage-centered point of view, by which in this case I refer to the creation of place-names, most importantly of hybrid ones

of any sort. Still, I consider it important to briefly discuss some questions related to the nature of the onomasticon.

In the light of the previous discussion about the original lexical meanings of names, I believe that it is not correct to assume that common nouns fulfill the role of categorization and establishing group membership, while proper nouns, or names, fulfill the role of identification, and singling out and labeling individual entities, places or persons. It should also be borne in mind that the onomasticon is not merely a list of names that occur in a language, but it is, according to the definition of Colman (2014: 1), the “repository for lexical information about names” which parallels with the lexicon that contains similar information about common words (Colman 2014: 2, 4). It should also be kept in mind, however, that there are some crucial differences between lexical items and proper nouns in terms of their developmental paths, behavior and usage patterns, and attention should be paid to whether we are dealing with them synchronically or diachronically in order to avoid conflicting the temporal dimensions, and one should also be wary not to conflate properties of the two groups as it can lead to gross misconceptions and misapprehensions. In the case of toponyms (whether hybrid or monolingual) it is especially important to take language change and the diachronic dimension into consideration. For the analysis of toponyms, Nicolaisen (1980) introduced the notion of onomastic dialects, which refers to those name-forming elements (generics and other formants) that inhabitants of a certain region typically use for building toponyms. Onomastic dialects can vary greatly from region to region, and a generic that is very frequent in one area might be completely absent from place-names of a different region. In a similar way as isoglosses define dialect boundaries, so-called isonyms can be established that define the boundaries of onomastic dialects, however, Nicolaisen (1980: 42) calls attention to the fact that the two are not exactly the same, but the underlying principle is still similar.

There exists a very strong link between the lexicon and the onomasticon, namely that every item of the onomasticon ultimately originates in the lexicon, and no new proper name can be coined without the language user being aware of the lexical content of the name (at least at the moment of its creation). Therefore, the onomasticon and the system of toponomastic elements is defined and determined by the system of the language and the system of the lexicon (cf. Ditrói 2017: 30). If we are talking about borrowed personal names or transferred place-names, then those items would have to originate in the source language’s lexicon and would have little or no connection to the receiving language’s lexicon. In such cases the borrowing is carried out only between the onomasticons of the two languages. In

those cases when place-name formants and generics are borrowed, a lexical link still has to exist because otherwise these elements would be functionless and speakers would not be able to use them productively¹⁶.

In keeping with the thread outlined above, for the purposes of the present dissertation and the analysis contained within, the lexicon and the onomasticon will not be sharply separated from each other and will not be treated as sharply distinguishable entities. With this in mind, then, it can be said that the generic elements of place-names are always appellatives and their purpose is to classify, doing so by categorizing the locality in question as an outlying farmstead, a secondary settlement or a village, just to name the most frequently occurring ones. Affixes build words, generics build toponyms.

In my view, place-names are best considered to be more akin to compounds than to derivatives, because they are created by conjoining two independent words, both of which have lexical meanings (and not grammatical meanings) and neither of which are grammaticalized, and neither of which are bound morphemes. A compound can be defined as a lexical unit that consists of more than one word which form a lexeme (Kunter 2011: 5), and in other words compounding entails the “lexicalization as a single unit, of syntactic sequences of independent words” (Colman 2014: 1). Place-names fulfill these requirements given that they consist of a sequence of independent words¹⁷ and are mostly created by word-formational processes such as derivation and compounding (Bába 2013: 107), with the most frequent compound patterns being Noun + Noun (as in *Asby* ‘village where ash-trees grow’ < ON *askr* ‘ash’ + *bý* ‘village’) and Adjective + Noun (as in *Bradford* ‘(place at) the broad ford’ < OE *brād* ‘broad’ + *ford*). This means that in general, two major categories of place-names can be set up based on their constituents: nominal compounds and attributive compounds.

From the diachronic perspective, though, affixes come about through the process of grammaticalization and they start out as free words which undergo the process of semantic bleaching before acquiring their final status as a derivational affix. Generics in toponyms also begin life as free words and they are also prone to losing their lexical meaning, however they do not become grammaticalized, and place-name formants and generics can only create new toponyms while they still have their lexical meaning, and hence are productive. Once the lexical words that are used as toponymic formants lose their original meaning or become

¹⁶ Although, it seems that analogical formations do not necessarily require this lexical link to exist, as speakers can conceive of the place-name elements that they use analogically as if they were formants or formative elements.

¹⁷ With the obvious exception of simplex names which contain only one element and are created through derivation.

extinct in the language, they become unproductive and incapable of creating new names. Affixes, on the other hand, acquire their grammatical meaning and function once they have lost their lexical meaning and have become grammaticalized, and when a word-formational rule becomes unproductive, the affixes fall out of use. This means that, similarly to compounds, derivatives can also get obscured and sometimes the originally complex word ceases to be analyzable to present-day speakers.

One half of the main argument here then is that if a free word that is used in a compound drops out of the language then speakers will not be able to create any more compound words with that item¹⁸, and if the free equivalent of a toponymic formant becomes extinct in the language then as a direct consequence, similarly as in the case of compounds, speakers will not be able to use that word to create new formations due to the lack of access to its meaning. This can be construed as a similarity between toponyms and compounds. The other half of the argument is that affixes behave in a different manner, namely that they do not carry and do not have to carry lexical meaning in order to be productive and speakers do not have to be aware of this. Herein lies a fundamental difference between compounds (in general and as toponyms) and affixes: constituents of compounds and generics in place-names have to carry lexical meaning in order to allow speakers to use them productively and creatively, whereas affixes do not, but they have to carry grammatical meaning and have to be in currency in any given stage of the language's historical development.

Bába (2013: 103-104) considers those elements of language to be name-formants with which speakers can create new names on the basis of already existing ones, and distinguishes between primary and secondary place-name formants. Primary formants are those elements of toponyms whose meaning in the place-name coincides with their lexical meaning, and secondary formants are those elements in which the lexical and toponymic meanings are separated and independent from each other. This distinction introduces an unnecessary extra label and thus an unnecessary extra level of complexity to the analysis of toponyms, because diachronically all toponymic elements' meaning coincides with their lexical meaning and through the semantic fossilization of the elements in place-names, their meanings can diverge from the lexical senses. Therefore, while on one level the distinction is justified, as fine-grained analyses are very important for uncovering the details of place-name formation, it will not be utilized in this dissertation due to the reasons outline above.

¹⁸ Unless of course they create them as analogical formations on the basis of the already fossilized (but to some degree still analyzable, even if by only folk etymological means) compounds, which is most frequently done for the purposes of word-play or to achieve a humorous effect.

In this dissertation it is assumed that basically appellative toponyms are right-headed endocentric compounds, which get their main category (i.e. ‘farmstead’, ‘hill’, ‘house’ etc.) from the lexical meaning of the generic. There are certain place-names, though, which were created via the application of derivational or inflectional affixes and these names typically lack a generic and in many cases are to be understood as elliptical expressions, for instance *Bolam* (Durham, Northumberland) ‘(place at) the tree-trunks’ < OE *bol* or ON *bolr* in the dat. pl. *bolum*. On the formal level, compound place-names consisting of a specific and a generic develop and behave in much the same way as non-transparent compounds do¹⁹. The main argument here, then is that a strict, binary differentiation of the lexicon and the onomasticon is unpractical, because names originate in the lexicon and elements of the onomasticon can also be used in a lexical function.

The names that speakers create have to be accepted by the speech community both linguistically and culturally. This means that they have to be formally and semantically transparent, they have to fit into the onomastic system of the language and they have to be created according to the accepted name-giving conventions and rules of that community (cf. Ditrói 2017: 25). It is crucial that the new names fit into the already existing system, and in the case of hybrid names of any sort, speakers of the involved languages have to be aware of each other’s naming patterns and also have to possess at least some degree of mutual intelligibility or bilingualism to be able to, on the one hand, create the names and on the other hand to make sense of them. The hybrid names created by the speakers either through cognate substitution or as new hybrid coinages should also conform to and mirror the way in which monolingual names are created. When place-name formation is involved in a contact situation, an awareness or knowledge of the other language’s onomasticon and naming conventions is also a prerequisite.

Generally speaking, already existing patterns and the analogical force they can exert on new coinages and on reanalysis play an important role in the formation and judgment of well-formed lexemes. The morphological competence of the speakers of a language allows them to create and recognize grammatical and well-formed lexical items by relying on the productive word-formational rules of their language. The same process is observable concerning the onomasticon. Onomastic models and patterns are of crucial importance, as they are required for speakers to acquire name-formational rules, and to allow them to

¹⁹ In this case the term non-transparent does not refer exclusively to exocentric compounds, which are by nature non-transparent, but rather to a specific kind of compounds, sometimes labeled ‘darkened compounds’ (see Section 3.2. below for a more detailed discussion of the nature of such lexical elements).

recognize lexical items as place-names (Ditrói 2017: 15). This means that both non-proper words and place-names have to be recognizable by speakers, because otherwise they would not be able to create new names and other speakers would not be able to recognize the new coinages. The relevance that this bears on the topic of the dissertation is that language users would not have been able to create and properly perceive and recognize hybrid names (especially the English-Scandinavian ones) if they had not been familiar with the other language's name-formational rules and had not been able to use them productively.

On the formal-structural level, as it has been hinted at before, place-names can be construed as compound words, which have been rendered opaque during the historical development of the language and had become darkened compounds by the time the language reached its current state. This obscuration can be of various types: phonetic-phonological, morphological, orthographic, or semantic, or any combination of these. In some extreme cases the process may result in all four features being affected and the originally polysyllabic word can be reduced to a monosyllabic one and all connection to its original constituents and semantic motivation will be completely lost. In the next sub-chapter it will be surveyed in what ways can place-names be affected by obscuration and what the reasons might be behind the present-day forms of toponyms.

3.2. Place-names as darkened compounds

A darkened (sometimes also referred to as obscured or opaque) compound can be defined as an originally transparent lexical unit, which comprises at least two independent lexical morphemes²⁰ and during the historical development of the given language at least one of these free morphemes has become “deconstructed [...] so that the meaning of the compound cannot be derived directly from the sum of its individual parts” (Bloomer 1999: 52²¹), or in extreme cases they have fused together into one single, monomorphemic lexicalized unit, no longer analyzable into separate words (cf. Čermák 2005: 35²²). For instance the well-known example of ModE *lord* < OE *hlaford* < *hlafweard* ‘guardian of the bread’ is considered to be a darkened compound which has become not only semantically

²⁰ Naturally, this does not exclude compounds which contain inflected or derived forms.

²¹ In his paper, Bloomer specifically surveys nominal compounds in English and German, but his theoretical points and findings can be extended to cover any type of obscured compound.

²² Čermák (2005) uses the term *obscured compounds* to refer to lexical units which have undergone both semantic and morpho-phonological obscuration. This understanding of the notion could be problematic though, concerning the definition of darkened compounds, as will be seen from the discussion of below. Therefore, I propose to use the term *obscured compound* to refer to any compound which has undergone any kind of obscuration.

deconstructed, but it has also contracted into a monomorphemic and monosyllabic unit. A compound can become opaque when it is no longer analyzable into its constituents (morphological and phonological obscuration) or when its meaning changes in such a way that the original motivation behind the emergence of the compound and its original semantic content is no longer accessible (semantic obscuration).

These two processes of phonological-morphological and semantic obscuration are not mutually exclusive, however, because either of them can separately affect transparent compounds or they can occur concurrently, as illustrated by the previously mentioned case of *lord*, in which the original meaning of the compound is changed and its original morphological and phonological form is not retained either. In a more strict, narrow sense only those are to be considered genuine darkened compounds which have been affected by morphological and phonological obscuration (regardless whether or not they were subject to semantic obscuration as well) because in those cases the lexeme in question appears to be a simplex one and its origin as a complex lexeme, comprising two (or more) elements has been obscured by phonological and morphological changes. On the other hand, those compounds which have only been obscured semantically are not genuine darkened compounds because they can still be segmented into their constituents, albeit their meaning has undergone change.

Semantic obscuration without morphological and phonological obscuration is typical of exocentric compounds, all of which have a meaning that is in some way external to the literal semantic content of the individual constituents, that is the meaning of the whole cannot be derived from the lexical meaning of the parts. These compounds, however, originate as semantically transparent ones, most of which are to be interpreted, from the moment they are coined, through synecdoche and metaphorical and metonymical figurative readings, and with time the original motivation behind the meaning extension is lost, which renders them semantically opaque. Cross-linguistically, including English, exocentric compounds have a very strong tendency to be an unproductive means of word-formation (Bauer 2008), meaning that they are very likely to be lexicalized, i.e. they “could no longer be produced according to productive rules” (Bauer 1988: 246). The following types of present-day exocentric compounds might be differentiated (based on Bauer 2008):

- (1) compounds which were originally **transparent**, **endocentric** and **non-figurative**, but their constituents have undergone semantic change and/or were lost from the lexicon but remained fossilized in the compound without accompanying phonological changes,

- (2) compounds which were originally **transparent**, **endocentric** and **metaphorical** but their constituents have undergone semantic change/demotivation and/or fell out of use in the language at large but remained fossilized in the compound without accompanying phonological changes, e.g. *blackmail*, which comes from *black* and Middle English *male* ‘rent, tribute’, referring to protection money paid by the English to the Scottish guards,
- (3) compounds which were originally **transparent**, **endocentric** and **synecdochical** (always referring to a person and sometimes also labeled *bahuvrihi*, from Sanskrit, meaning ‘much rice’, i.e. someone who is rich) but their constituents have undergone semantic change/demotivation and/or fell out of use in the language at large but remained fossilized in the compound without phonological changes, e.g. *redcap*, which refers to American railway porters in a pars pro toto relationship, because of the red cap they traditionally wore.

If we take headedness and hyponymy – as Bloomfield (1935: 235-236) originally proposed²³ – to be the main determiners of endocentricity versus exocentricity, then the synchronic point of view mentioned at the beginning of the list means that looking at constructions like these from the perspective of present day English and present day English usage, they seem to conform to the criteria of exocentricity, i.e. they do not have a head and the whole compound is not the hyponym of the rightmost element. However, as Bauer (2008) also notes, if they are looked at in a historical linguistic context, it emerges, as could be seen from the list, that actually none of them are truly and diachronically exocentric because they originate as endocentric compounds but have been (i) figurative since the beginning, (ii) acquired a figurative interpretation later on or (iii) intervenient historical changes (semantic and/or lexical, as described in the list) have rendered them opaque. The categories outlined in the list above seem to be united by the fact that their members are **unproductive** and therefore necessarily **lexicalized**. The unproductive and lexicalized nature of these compounds are brought about by morphological and phonological changes and/or fossilization of otherwise lost lexical items (Brinton & Traugott 2005: 50).

Bloomer (1999) presents an analysis of twelve possible stages of obscurity ranging from instances of compounds in which both constituents are formally and semantically fully

²³ Headedness and hyponymy in the case of compounds refers to whether or not they possess a semantic head (if they do then they are endocentric, if they do not then they are exocentric), and whether or not the rightmost element, which is typically the head, is the hyperonym of the entire compound (if it is then the compound is endocentric, if it is not, then it is exocentric).

transparent to those which have been rendered formally and semantically fully opaque and have even become monosyllabic. His taxonomy and argumentation will be adopted in this dissertation for the analysis of various place-name formations and the discussion of their developmental background in this chapter. Bloomer's main argument (1999: 57-61) hinges on the notion of isolation, whereby the constituents of the compound word become semantically divorced from their 'free' counterparts due to the fact that the reference of compounds is much more restricted and singular than that of free words. This seems to echo Clark's (1992a: 453) observation about place-names, in that once the so-called onomastic divorce sets in, they "draw partly aloof from the language at large" which can also be said about compounded lexical elements, although that divorce will not be of an onomastic nature but the underlying principle would be the same. Given enough time this semantic isolation is then coupled with divergence resulting most frequently from the phonological changes of stress (re)assignment and assimilation, and the disappearance of one or both of the words constituting the compound from the lexicon²⁴ of the language while they are still preserved in a fossilized form in the compound.

If we apply this reasoning to place-names it can easily be seen from the discussion above that the exact same processes take place in toponyms too. Given that place-names are compounds in the first place and that they have only one referent, their constituents are semantically isolated from the free words and when they diverge formally and/or the word falls out of use in the language at large, toponyms step on the same path of development as 'regular' compounds. Once the compound or place-name is seen as referring to a singular entity and as having a singular meaning, it will be more prone to formal and structural changes and might be affected by them more severely and more thoroughly than free lexical items, or in some cases certain grammatical features (such as the OE genitive *-n* or the Scandinavian *-ar* genitive²⁵) might get fossilized in toponyms (cf. Clark 1992a: 485, Coates 2006a: 314).

If we turn our attention more to toponyms, it can be observed that from the present-day perspective they are also obscured and become, similarly to exocentric and especially

²⁴ The most frequent scenarios for loss of lexical items include socio-cultural development, language contact situations resulting in loanwords entering the language and ousting native terms, competition among synonyms, cases of homonymic clash, and semantic changes such as specialization or narrowing which might lead to certain lexical items to be used in very restricted ways and only in specific contexts and this might even lead to them disappearing from the language altogether.

²⁵ For instance *Coventry* (West Midlands) 'tree of a man called **Cofa*' < OE pers. n. **Cofa* + genitive *-n* + OE *trēow* 'tree'; or *Litherland* (Merseyside) 'cultivated land on a slope' < ON *hlīðar* 'slope (gen.)' + OE/ON *land*. The importance and implications of the preservation of Old Norse inflectional endings in hybrid names will be elaborated on in a later subchapter.

darkened compounds, unproductive and lexicalized. Furthermore, they are one step beyond exocentric compounds, because place-names can in fact be construed as darkened compounds which are affected both by semantic and morphological-phonological obscuration, possibly in that order, given that the “semantic divorce from common vocabulary lays name-material especially open to phonological change” (Clark 1992a: 485). As discussed in the previous sections, this semantic divorce happens through the processes of onymisation, isolation, and onomastic divorce. Toponyms and elements in them are detached from their original lexical meanings as the lexical items in question gradually fall out of use from the language. Place-names are affected by lexicalization, coalescence, morpho-phonological fusion, demotivation and semantic obscuration in much the same way as darkened compounds are. The difference, though, between a place-name and a darkened compound lies in the mono-referential nature of place-names which means that they refer to only one single referent.

Instead of binary oppositions and distinctions, it is perhaps more advisable to place lexical morphemes, endocentric compounds, exocentric compounds, darkened compounds and place-names on a continuum, as I propose below in *Figure 2*:

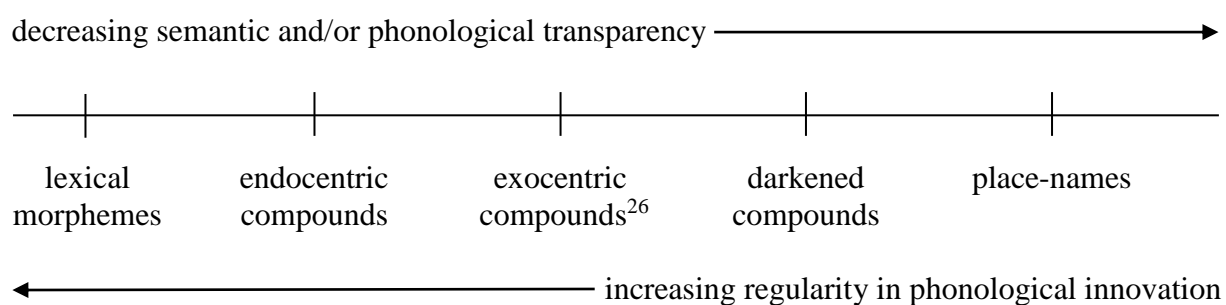


Figure 2. The continuum of compounds

Going left to right along the continuum, from a synchronic perspective lexical morphemes and endocentric compounds have the highest degree of semantic and morphological-phonological transparency, while darkened compounds and place-names have the lowest. Members of the continuum also exhibit increasing degrees of lexicalization when going from the left edge towards the right edge. Going from right to left we can observe that members of the continuum show increasing resistance to phonological change, meaning that

²⁶ The label *exocentric compounds* is used in this figure only for convenience, and it comprises those compounds which have undergone any sort of semantic change so that their original endocentric nature has become obscured, as well as those which are originally metaphoric or metonymic but otherwise regular endocentric compounds, however their meaning is non-compositional and therefore conform to Bloomfield’s (1935: 235-236) classic definition of exocentricity. Those lexical items which have undergone both semantic and morpho-phonological changes are to be considered darkened compounds.

lexical morphemes undergo regular, systematic phonological changes (such as the Great Vowel Shift, for instance) while darkened compounds and place-names seem to be subject to a wider variety of changes which appear to affect them more sporadically and less systematically than major sound changes affect lexical morphemes.

In conclusion, processes of obscuration affect place-names in the same way as they affect lexical compounds: the onomastic divorce in place-names detaches the semantic properties of name elements from their free, lexical counterparts, and in lexical compounds isolation and restriction of reference paves the way for obscuration. Compounding was a highly productive means of word-formation in Old English, which saw a marked decline in the Middle English period, when a large number of OE compounds were completely replaced by French words while others underwent obscuration. Just as the Proto-Indo-European word-formational and tense-marking process of root vowel gradation obsolesced and became lexicalized, so were the Old English compounding patterns and compounds abandoned with the surviving ones becoming darkened and lexicalized. Due to the rather hectic nature of sound changes which affected place-names, clear explanation of historical developments in toponyms is not possible, as the next subchapter will show.

3.3. Sound change and morphological change in place-names

Discussion of the phonological and morphological changes and obscuration processes affecting place-names have not received nearly as much scholarly attention in the available literature as their socio-historical context, their linguistic properties in general, and their relevance as linguistic items. Papers and monographs discussing place-names seem to largely disregard the historical changes that they can become subject to. In this section only a cursory exploration will be presented, based on some preliminary observations and some scholarly articles. A full treatment of the issue would only be possible after thorough inspection of the changes, which unfortunately is not the focal point of this dissertation.

In what is perhaps the first systematic treatment of the issue, Mutschmann (1913) presents an overview of the phonological changes affecting the toponyms of Nottinghamshire (Mutschmann 1913: 159-163). Some of the major general developmental tendencies he observes are the following:

- (i) shortening of long vowels in stressed syllables before certain consonant clusters, although the exact clusters initiating this change are not described,

- (ii) monophthongization of $\bar{e}o$ into e and $\bar{e}a$ into a in the same environment as the previous point,²⁷
- (iii) loss of the medial consonant in three-consonant clusters,
- (iv) assimilation of the first consonant to the second in two-consonant clusters, and
- (v) glide insertion²⁸.

The changes outlined here – with the exception of the final one – all fall within the categories of lenition, reduction, simplification, and loss. Such changes of reduction are most probably the result of weak stress.

In accordance with the previous discussion on the semantic characteristics of place-names, Clark (1991: 284) claims that the “aberrant” phonological development in names is due to their special semantic character. What this means is that due to the onomastic divorce and the loss of the semantic link with free lexical items, toponyms become especially prone to linguistic changes, such as “assimilation or dissimilation, elision and syncope, procliticization, folk-etymology” (ibid.) and other processes such as the shortening and monophthongization as described by Mutschmann (1913) above, or the insertion of inetymological epenthetic vowels. Perhaps the most typical phonetic change in place-names is the reduction of vowels, especially in weakly stressed syllables and in the case of medial syllables this often results in their syncope²⁹.

On the whole, it can be observed that in the case of place-names general native tendencies that operate in free lexical items operate in a less restricted manner and can affect place-name elements sooner than their free equivalents (cf. Colman 1992: 61), which is then amplified by the general semantic properties of the toponyms and by the loss of the free equivalents, or common word cognates in Colman’s terminology, of toponymic elements. Clark (1991: 284-287) assumes that names (both toponyms and personal names) are affected by a different kind of change than free lexical elements, albeit in principle the processes are the same for which she uses the term “onomastic sound-change.” As we have seen, this is a

²⁷ Mutschmann only provides a couple of examples for each change he mentions, but fails to indicate exactly what consonant clusters and which positions bring about these changes. Mutschmann also treats monophthongization under the heading of vowel shortening, I separated the two phenomena in my description.

²⁸ For the full list of changes with examples see Mutschmann (1913: 159-163).

²⁹ Certain place-name elements, such as OE *ceaster* ‘Roman camp, fort’ can function as generics, specifics, or simplex names and can develop differently depending on which function they appear in. If *ceaster* occurs as the generic then due to its weakly stressed word-final position it is prone to vowel reduction, syncope, and apocope. On the other hand if it occurs as the specific or as a simplex name then its is almost uniformly pronounced as /tʃestə/ (cf. Clark 1991: 286). This proves that the sound changes affecting toponyms are not random but have their internal consistency and are conditioned on the morphophonological structure of the word.

valid assumption as evidenced by the examples given above, and would imply that the semantic properties of words play a significant role in the way in which they are affected by sound change and morphological change. However, such an idea must be approached with care, as Clark also raises awareness about the pitfall of assuming that non-standard sound changes take place in toponyms, as it can lead to “explaining away” any deviation from normal development instead of careful analysis.

In an analysis of the inscriptions on Anglo-Saxon coins, Colman (1992: 59-67) discusses phonological changes in Old English personal names that are not attested in related common words, and, similarly to Clark (1991), also concludes that there are certain sound changes that only affect names and not common nouns. Following Clark (1991) and Colman (1992), Coates (2006b) also refers to such changes as “onomastic sound-change” and points out that the importance of a separate, or at least somewhat different, mechanism of change such as this lies in the implication that apart from system-internal factors, sound change can also be conditioned by non-phonetic phenomena (i.e. semantic features) and by sociolinguistic constraints. As a result, Coates (2006b: 265) makes the claim that “onomastic sound-changes need not themselves be regular” which we can assume to mean that phonetic changes affecting toponyms can be entirely random. I do not necessarily agree with his claim, as such an assertion would warrant and even necessitate further research into this topic in order to provide a conclusive proof.

Apart from onomastic sound changes, regular sound changes also affect place-names. In accordance with the previous two subchapters, my stance is that place-names are to be classified as nouns, and in my view as compounds. Productively coined endocentric compounds are subject to the compound stress rule, which means that the primary stress of compounds must fall on the first constituent and the second must be weakly stressed or completely unstressed. This can be observed in the present-day pronunciation of county names, whose final and penultimate syllables are usually zero-stressed and reduced to schwa, for instance *Yorkshire* /jɔ:kʃə/, *Lancashire* /læŋkəʃə/, or in names of weekdays such as *Tuesday* /t(j)u:zdi/ < OE *tiwesdæg*. In the case of *Tuesday*, phonological reduction and obscuration took place which led to the emergence of a darkened compound, meaning Tiw’s day with a reduced second element that is still analyzable semantically plus an obscured first element, which is a personal name no longer in currency. The structure is similar to that of *Grimston*, Grimr’s village, for example, in which the second element is OE *tūn* which in non-toponymic usage regularly yields ModE *town* due to the Great Vowel Shift affecting the long vowel in the word, but in place-names it is invariably pronounced as /tən/ in ModE because it

shortened prior to the operation of the Great Vowel Shift and therefore did not participate in the change.

System-internal and semantically conditioned onomastic sound changes are prime candidates for the explanation of the historical development of toponyms. However, it has long been proposed that the Domesday Book's orthography and other scribal influences could also have played a role in shaping the form of English place-names. Through a series of papers, Clark (1991, 1992a, 1992b, 1992c) convincingly demonstrated that attributing 'non-standard' spellings and phonological features of certain English place-names to the alleged sporadic and irregular 'Anglo-Norman scribal influence' on the forms in question is an untenable hypothesis. The core of her argumentation rests on the fact that the documents from which names are extracted for onomastic study are chiefly formulaic administrative records and all of the documents in which 'deviant' or 'irregular' spellings are found *are* administrative records and such forms should never be examined as if they existed in a vacuum, but the then current orthographic conventions must always be factored in (Clark 1991: 276, 1992a: 455, 1992b: 124). Furthermore, she asserts that native speaking of Norman French could never have been widespread outside the higher levels of society and the Domesday Book's primary scribe could be identified as "English-trained and probably English-speaking" (Clark 1992b: 122-123). Cameron (1996: 91) accepts these views of questioning the significance of scribal influence, albeit rather cautiously, while maintaining earlier that Domesday Book scribes "represent[ed] English sounds by the nearest equivalent in their own language"³⁰ (Cameron 1996: 16). This, however, in itself may not be categorically deemed false because in essence the orthographical transformation of English place-names in this fashion was necessitated by the fact that they had to be accommodated in the framework of a non-English, foreign document, written in a language that was not English.

Clark's reasoning in the articles cited above, the social history of the post-Conquest period and common sense would dictate then that it is erroneous to attribute non-standard³¹ spellings and even seemingly unique sound changes observable in English toponyms solely to scribal influence. The phenomena of morphological, phonological and semantic obscuration discussed in the previous subchapter could perhaps prove to be a better candidate. As Čermák (2005: 39) also observes, the transition from Late Old English to Middle English "appears to

³⁰ However, from the present's perspective we cannot tell what were the exact sound values of these graphemic representations and what letter-to-sound rules obtained.

³¹ The concept of a 'standard' in the Old English period, however, is a rather dodgy one, given that there existed no regularized, standardized variety with a regularized, standardized spelling.

have been a particularly favourable time for obscuration processes in compound words” which is exactly due to the significant phonological, morphological and syntactic (the loss of grammatical gender, the acceleration of the already ongoing gradual disintegration of the inflectional system and the accompanying syncretism, and the consequent loss of the relatively free word order and the language’s typological shift towards the morpho-syntactically analytic category), socio-historical (the Norman Conquest and its effects on the upper layers of society) and sociolinguistic (the superstratal influence of Norman French on Old and Middle English) changes taking place during this period.

As it has also been mentioned previously, place-names seem to be especially prone to morphological and phonological change, which is related to their semantic characteristics. It should also be noted though that, presumably during the above mentioned transitional period, many of the specifics and generics that are found in modern English place-names have been lost from the language, which could have also contributed to their semantic obscuration (if not initiated it), which then paved the way for morphological and phonological obscuration. Therefore, scribal influence and onymisation cannot really be held responsible for the obscuration of place-names, but the loss of certain lexical items found in toponyms and the linguistic, historical and sociolinguistic changes occurring during the transition from Old to Middle English could provide a more reasonable explanation, as it also does for darkened compounds. Such an explanation would entail that the significant changes in linguistic and lexical structure outlined above meant that many elements in place-names became fossilized and when this was accompanied by semantic isolation and monoreferentiality the process of obscuration sped up. Furthermore, foreign influence affected the English language holistically more significantly than it affected the names specifically, including their spelling, pronunciation and obscuration, so it cannot be said that these processes targeted only toponyms and anthroponyms.

3.4. The genesis of hybrid toponyms

As it was discussed in Chapter II, the scholarly literature does not have a lot to say about the nature and development of hybrid place-names, let alone their relationship to code-switching. In the present chapter the genesis and general linguistic features of place-names were introduced, and as a conclusion for Chapter III the possible ways in which hybrid toponyms can come about will be briefly outlined. After the empirical analysis is conducted,

the results will be compared with the types described in this section to see how they are represented in the corpus of place-names.

In reference to Section 2.1 we can distinguish between the following types of hybrid place-names if we adhere to the broad definition of hybrid toponyms:

- (i) bilingual hybridization, i.e. originally hybrid names, like presumably most of the appellative ON-OE names that were created as originally hybrid toponyms and are the best candidates for the utilization of code-mixing,
- (ii) hybridization via personal name specifics, in these cases the speaker does not really have any other option but to create a hybrid name because foreign personal names are typically used only in the form that they occur in the other language,
- (iii) hybridization via sound change, i.e. the so-called Scandinavianized names which were created by substituting those OE sounds with ON ones that were not available in the Scandinavian speakers' phonetic inventory,
- (iv) hybridization via cognate substitution, which are most prominently observable in English-Scandinavian names, especially so in presumably early coinages created after the initial wave of settlement,
- (v) hybridization via unproductive borrowing, e.g. *Yorkshire*, *Devonshire* in which cases the first element cannot be used in a productive manner because they are synchronically unanalyzable already existing place-names or river-names (mostly Celtic ones) therefore chiefly analogical formations belong here, this type typically involves the specific being of foreign origin (cf. Minkova & Stockwell 2009: 36),
- (vi) as a subcategory of (v) above, those hybrid names also belong here which were created by superadding a generic of the invading language onto the already existing names of the previous population (for instance the expansion of Celtic names after the Germanic influx) which often results in the emergence of tautological names (cf. Bölcskei 2012: 158-159 on Celtic-English hybrid toponyms created this way),
- (vii) hybridization via productive borrowing, which are the most frequently occurring ON-OE hybrids, especially with generics like *bý*, *þorp*, and *tún* that can be found in the highest numbers on OE-ON and ON-OE hybrid names, respectively.

In a paper appearing in the same volume of conference proceedings (Andersson, Brylla & Rostvik 1980) as Walther's contribution above hybrid names cited in Section 2.1, Sköld (1980) provides a discussion of place-names in relation to what he calls "code shifting" ("kodväxling"). His account is not so much focused on code-switching as such, but on the use of exonymic equivalents of place-names in various languages and on the occurrence of different names for the same settlement in multilingual areas, based on a brief personal observation of Lappish place-names being translated into Swedish by Lappish speakers. In his paper, Sköld (1980: 268-269) also proposes a four-way categorization of the use of foreign toponyms in a given language and the use of toponyms in a multilingual context. The categories he sets up are the following: quotation loans, foreign names, homely names, and translated names. Quotation loans refer to those foreign place-names which are used in their original form without any change or integration into the recipient language, foreign names are ones which have undergone phonological adaptation and have been rendered more congruent with the receiving language's pronunciation, the category of homely names should perhaps be better called nativized names, as these are the ones which have been fully integrated into the language, and translated names straightforwardly refer to names which have been calqued into the recipient language.

In my view, quotation loans and homely names can be conceived of as representing the two extreme ends of a scale ranging from no integration and adaptation whatsoever (quotation loans) to complete integration (homely names), with fuzzy boundaries in between to accommodate for the occurrence of uncertain instances. This categorization is a useful starting point for the analysis of the way in which foreign names are treated in a language. In prolonged language contact scenarios involving settlement most of the names start out as quotation loans and in due course become nativized names with often an accompanying obscuration of their original meaning and structure. Concerning hybrid names, however, only those elements can be used for forming toponyms which match the phonological system of the recipient language, either through integration (in which case the elements are likely to have spent some time in the system of the other language) or by the foreign element already being congruent with the receiving language's system.

In certain cases, instead of being the product of a bilingual mixed community, hybridization of place-names can come about through the borrowing of generics from one language into the other. When this happens, familiarity with the lexical content of the borrowed generics is also vital, otherwise the entire act of borrowing would be pointless, since any act of borrowing requires that speakers of the receiving language be aware of the

meaning of the foreign elements. However, those place-name formations which are created via the utilization of borrowed words cannot be construed as genuine hybrid toponyms, because they were not created by employing elements of two languages, but by employing a native element and an originally foreign one which has been taken over and integrated into the receiving language. Obviously, as with any borrowing, the issue of the time depth of the borrowing has to be factored in too, and the problem of distinguishing between historical code-switching and borrowing is also relevant, as discussed in Section 2.3.

Without exception, the hybrid place-names, whatever their origin is, will be subject to the regular morphological changes and sound changes (and onomastic sound changes) of the language in which they are used, and some of these developments will be conditioned and constrained by semantic character of the names. Besides the regular sound changes resulting from the historical development of the language at large and the onomastic sound changes resulting from the special semantic features of names, analogical reformation, i.e. folk etymology can also contribute to the emergence of the present-day form of toponyms.

IV. Layers of Language and Settlement in Britain

This chapter serves as a transition and bridge between the theory-oriented second and third chapter and the actual empirical analysis of the next chapter. Here, the historical context in which hybrid toponyms emerged will be surveyed very briefly, with a more pronounced focus on the relationship of the languages spoken in given time periods. The main focus of the chapter is not so much to provide a historical account of events, but to provide a discussion of the stratal relationships of the languages embedded in a historical context. The typical place-name formations, specifics, and generics for each layer that is introduced here will be described and discussed in detail in the next chapter.

4.1. Defining the broad historical context

The roots and origins of whom we would today consider English people reach back several thousands of years to the European continent, to present-day Denmark and southern Sweden. Although people have inhabited this area from 10,000 B.C. onwards, the first tangible traces of Germanic tribes, in the writings of Greek and Roman historians, are only as recent as 200 B.C., with the first written evidence created by the Germanic tribes themselves (the runic inscriptions) dating from around A.D. 150. However, this does not necessarily mean, that there were no Germanic tribes in this region prior to 200 B.C. Presumably around the third millennium before Christ, the so-called Battle-ax Culture (also known as the Corded Ware Culture), who are believed to be related to the Indo-Europeans, and with whom Germanic tribes are usually associated, settled in this area (Robinson 1992: 16).

While on the continent the Jutland peninsula and the surrounding areas were inhabited by Germanic peoples, the British Isles gave home to a number of different tribes, mostly of Celtic origin, among whom we find the Britons, as well as the Gaelic Scots, who presumably migrated to the island from the Central and Western European regions. Archaeological evidence has been found that supports the presumption that Germanic Angles inhabited regions on the southeastern coast of the island before the major wave of Germanic invasion of the mid-fifth century. Prior to the Anglo-Saxon takeover of the island, the Celtic population was subjugated by the Romans in AD 43, when Claudius annexed the island to the Roman Empire, under the name Britannia. Roman rule lasted until 410, when the collapsing Roman Empire had to withdraw from peripheral provinces such as Britannia, after which the Celtic tribes had to suffer attacks from the tribes living in the northern territories, namely the Scots

and the Picts (Jones 1990: 45). Since the Celts had been weakened under the foreign rule, they could not defend themselves from these attacks, so they decided to invite Saxon mercenaries from the east for help.

Soon after the Romans left, the Saxon tribes, led by the legendary Hengist and Horsa arrived at the island, who have then successfully defeated the raiding Scots and Picts. This, however, meant that the Saxons, who allegedly came to protect the Britons, turned against them and set out to conquer their land, at which they succeeded. Soon masses of people were brought over to the island from the Germanic homeland. Thus, over a century the Anglo-Saxons managed to take over the island and set up the Heptarchy, that is the seven kingdoms; while the indigenous population of Britons were forced back into Wales and Cornwall (Rot 1992: 95).

Towards the end of the 8th century England was raided and invaded by Scandinavians in two distinct waves, who established the Danelaw in the north-eastern part of the island, and their presence lasted until the Norman Conquest of 1066. The Scandinavians also settled in large numbers and exerted significant influence on the place-names of Britain and also on the English language. In the aftermath of the Norman Conquest, the English aristocracy was replaced by a Norman French ruling class, and even though Normans did not settle in large numbers in England and did not leave such a mark on toponyms as Norse settlers did, the influence of Norman French was also significant on English.

4.2. Pre-Celtic and Celtic Britain

As it was established in the introductory chapter of the dissertation, Britain was subject to a number of waves of migration and conquest, which meant that the languages spoken by the different groups came into contact with each other. Unfortunately, we do not have a lot of linguistic evidence of pre-Celtic populations apart from some speculated remnant forms in place-names and river-names that have been carried over into Celtic place-names, as the Celtic settlers formed a superstrate over the pre-Iron Age population. Bölcskei (2012: 153) assumes that the pre-Celtic indigenous population of Britain was mostly of Iberian origin. The pre-Celtic layer of English toponyms is associated with the Old European substrate of populations and languages, that according to some hypotheses (e.g. Vennemann 1994) existed before the Indo-Europeans expanded westwards into Europe and are pre-IE in origin. However, the mainstream view (e.g. Kitson 1996) still holds that these Old European toponyms and hydronyms are in fact of an Indo-European ancestry.

It is difficult to date when exactly the Celtic tribes could have arrived in Britain, with estimates ranging from as early as 2000 BC to as late as 600 BC (Hickey 1995: 93), as Celts have been identified with a variety of Central European Bronze Age cultures (e.g. the Beaker Culture) and Iron Age cultures (e.g. the Hallstatt Culture and the La Tène Culture). A further complicating factor is that the Celtic tribes did not arrive in one large group but there were several strands and layers of Celt-related cultures and populations that settled in Britain over the course of centuries before the Roman subjugation. On the other hand what we do know perhaps without doubt is that the language spoken by the Celtic tribes was a substrate language in relation to (pre-)Old English after the arrival of the continental Germanic tribes. However, the exact nature of the relationship between the two peoples and the influence the Celtic languages exerted on English can be subject to debate.

What is certain is that Celtic influence on English, as Vennemann (2011: 226) also confirms, is the typical example of substratal influence, which was actuated when Celtic speakers shifted to English. The Celtic language, being a substrate to Old English, exerted limited lexical influence but it had more significant structural effects on English (cf. Tristram 2007, Filppula 2008, Vennemann 2011: 226-230), as it is usual in the context of dominantly substratal contact and shift (Hickey 2010a: 7). The Anglo-Saxons occupied the higher ranks of society, while the Celts were in a subordinate, subjugated position. A typical feature of substratal language shift is the appearance of structural (phonological, morphological, syntactic) and pragmatic changes in the superstrate language which in the case of Celtic-English contact can be observed, according to Vennemann (2011: 227-230), in for instance glide insertions, certain possessive constructions, the emergence of the *-ing* present participle in the southern varieties instead of the *-and(e)/-end(e)* suffix used in other dialects (cf. also Tristram 2007: 210-211), the emergence of *do*-support (cf. also Tristram 2007: 211-213, Filppula 2008), the retention of the interdental fricative which was lost in all Germanic languages with the exception of English and Icelandic, and on the pragmatic level in the way in which English forms answers to yes/no questions that is distinct from the other Germanic languages. To what extent features like these can in fact be attributed to Celtic substratal influence is subject to debate that will not be elaborated on here (see for instance Filppula 2008 and Filppula, Klemola & Paulasto 2008: 223-257 for a more in-depth discussion).

Since the linguistic relationship of Celtic and Old English was a substrate/superstrate one, the social relationship between their speakers could not have been a very different one either. A likely scenario was that the Anglo-Saxons occupied the upper layers of society while the Celts occupied the lower ranks. According to the widely held theory, also recapitulated

above following Rot (1992), the invading Germanic tribes drove the Celts to the fringes of the island and displaced them from their settlements. Hickey (1995: 103-104) questions this hypothesis, citing a lack of any possible incentive on behalf of the Germanic invaders to drive out the Celts, and, even more convincingly, the linguistic piece of evidence that is the word *wealh*, meaning ‘foreigner, Welshman,’ and later ‘slave’. This meaning of the word is taken to be indicative of the status of the Celtic population in relation to the Germanic tribes, and so is the wealth of place-names indicating settlements of Welsh and British inhabitants as well as their coexistence with the English population (cf. also Filppula 2008: 155-156, Bölskei 2011: 200, Schrijver 2014: 20).

In the Middle English period the inflectional system of the English language gradually disintegrated resulting in widespread case syncretism, and grammatical gender was also lost. Many scholars would ascribe the initiation of such changes solely to language contact situations, in which the Celtic, Norse, and French language have equally been identified as culprits. Bertacca (2007: 26-27) proposes that language contact with Celtic and later on with Norse and Norman French did not initiate the collapse of the English inflectional system, only accelerated the already present tendencies, and emphasizes that such a process is absolutely normal. Furthermore, Bertacca (ibid.) calls for the careful (re)assessment of the influence that Celtic (specifically Late British) could have exerted on Old English in the light of the available linguistic data and observations.

The view that the various contact situations did not induce the collapse of the English inflectional system but only contributed to it is still considered to be valid, because language-internal processes have proven to be better candidates for it. However, recently views that attempt to diminish Celtic influence and downplay its importance have been challenged, especially by Bertacca (2007), Filppula (2008), Filppula, Klemola & Paulasto (2008), and Vennemann (2011: 226-230) who goes as far as asserting that the shared innovations of the North Sea Germanic languages arose through contact with a “common North Sea Celtic substratum of the Insular Celtic variety” (given that Central European Celtic speakers are known to have been in contact with continental Germanic communities inhabiting areas north of the Celtic cultures) and that “standard English should be studied as one of the Celtic Englishes” owing to the structural influences.

In conclusion, the Celtic speakers first came into contact with the pre-Celtic population who supplied some place-name elements. Secondly, the Romans subjugated the Celts and did not tamper much with the Celtic toponyms they encountered. Thirdly, the Germanic tribes conquered Britain and formed a linguistic superstrate over Celtic that exerted

little lexical influence over Old English but caused more significant and observable structural and pragmatic changes in the superstrate. Concerning toponyms, some Celtic names and elements made their way into English and in some cases Celtic names were extended with the addition of an explanatory English generic, but the number of Brittonic place-names taken over by English is still low (cf. Coates 2007: 175).

4.3. Language in Roman Britain

The history of Roman Britain is traditionally considered to begin with Claudius in 43 AD when the island became part of the Roman Empire as the province Britannia³². The Celtic inhabitants of the island under Roman rule formed a substratum and were largely assimilated and incorporated by Rome, whereas those Celtic populations that lived outside the Roman boundaries of influence in Britain were left very much unaffected (Salway 2010: 1). The Celts were subdued and Romanized with client kings loyal to Rome ruling over them (Jones 1987: 128, Salway 2010: 16-17), the Latin language began to be used and widespread Latin-Celtic bilingualism is assumed to have existed (Schrijver 2007: 165). Changes related to civilization affecting the local Celtic population included urbanization, Christianization, and the introduction of literacy (Bölcskei 2012: 154), whereby Rome consciously aspired to secure their power and ideologically reshape the local population and make them less hostile towards Rome and the Latin language (Jones 1987: 127). The end of Roman Britain came with the gradual disintegration of the Roman Empire, as it had to give up and withdraw from peripheral provinces such as Britannia, thus Roman rule in Britain ended in 410 AD, after which the Germanic conquest commenced around the mid-5th century.

In contrast with the overwhelmingly oral culture of Celtic Britain, the society of Roman Britain was largely literate, while the Celtic language was still maintained with Latin being widely spoken even in late Roman Britain (Schrijver 2014: 15), and being the primary language of writing. Schrijver (2007: 165) establishes two zones of Roman influence in Britain: the northern and western “Highland Zone” and the heavily Romanized eastern “Lowland Zone.” Latin was spoken especially in the Lowland Zone where it was the predominant language, and where a Latinized variety of British Celtic was also in currency. Medieval (and thus modern) Welsh, Cornish, and Breton are the continuations of the less Latinized Highland British Celtic variety. Not only the geographical spread of Latin was

³² Although it should be noted that the picture was more complicated than this, because due to the resistance of Caratacus the Roman conquest nearly failed to succeed.

uneven, but the use of Latin also varied greatly with social status and occupation, as people involved in some way with the Roman military and Roman administration as well as those who lived in urban areas and owned estates were exposed more to Latin and were more frequent and proficient users of the language than those who lived in villages and were not involved with Roman affairs (Schrijver 2007: 165, cf. also Rivet & Smith 1982: 12).

Latin never fully displaced British Celtic, not even in the Romanized Lowland Zone, but it is very likely that speakers of Early British Celtic shifted to the use of Vulgar Latin, in which case it would have been a substratal language shift. Rivet & Smith (1982: 13-14) note that there is no written evidence of Brittonic being used in Roman times which would lead one to believe that British Celtic was displaced entirely by Latin, but in reality the reason is that the language of the Celts was not a written one therefore no written artefacts were produced. According to the prevailing scholarly opinion, by the 4th and 5th century AD in the Lowland Zone Latin must have been the main language of a large proportion of the population or even of the majority (Laker 2010: 6). Furthermore, (Vulgar) Latin did exert linguistic influence on British Celtic, as it was in a superstrate position in relation to British, but still very little is known about the variety of Latin that was in use in these Celtic areas (cf. Laker 2010: 117). Schrijver (2007) demonstrates that Highland and Lowland British Celtic developed along different paths and received different degrees of Latin influence, with Latin-induced innovations of a phonological and morphosyntactic nature only spreading into the Highland variety after the end of Roman rule as a result of the assimilation of the now low-prestige speakers of Late Latin in the Highland Zone.

Assessing the stratal relationship of British Celtic and Latin seems to be difficult at first because the extensive phonological influence of Vulgar Latin on Lowland British Celtic would point towards substratal effects on the indigenous language, but the dominance of Latin and the extensive lexical borrowing from Latin, especially in the Lowland Zone would indicate the superstratal status of the invaders' language. Schrijver (2002: 87-88) confirms that "early-first-millennium Latin" was a "politically, socio-economically, and culturally high-prestige" language and that it contributed a significant amount of loanwords to British Celtic, which is therefore indicative of a superstrate. The phonological changes brought about by Latin are the results of linguistic Romanization, which in the provinces typically resulted in the complete disappearance of the original language spoken in the area, with only three exceptions. Britain was the only western Roman province that lost the "Romance speech" (Jones 1987: 126), Basque and Albanian being the other two local languages that persisted, and the reason for the survival of these languages is that they were spoken in relatively remote

areas (Schrijver 2014: 31-32) and were large enough to resist complete Romanization. In Roman Britain, much like anywhere else in the Empire, local languages were in a subordinate position as compared to Latin which exerted pressure on their phonological, morphological, and syntactic subsystems, leading to the emergence of a sociolinguistic situation that is reminiscent of language death.

In conclusion, following the Roman subjugation, British Celtic (or Brittonic) split into two spheres of influence: the little-affected Highland variety and the heavily Romanized Lowland British Celtic variety. The influence of Latin can be observed in two phases: in a phase of expansion, under Roman rule, and in a phase of contraction, after the Roman withdrawal and the Germanic conquest (Schrijver 2002). The expansion phase is characterized by significant phonological changes and the extensive importation of Vulgar Latin loanwords into British Celtic as a result of Rome's conscious assimilation policy, and also substratal language shift from Brittonic to Latin took place. In the receding phase, Latin became the substrate to Celtic and the language of the Germanic invaders, with heavy lexical borrowing from Brittonic and (pre-)Old English, culminating in language death, and the discontinuation of the use of Latin at least in its late spoken form and in the function of an everyday language. Unstable bilingualism is postulated to have been present in both phases. In the Old English period, Latin had a high prestige as the language of the church and of education, with extensive borrowings from Latin into Old English in these domains.

4.4. Anglo-Saxon England

The story of the Anglo-Saxon settlement and conquest in Britain begins in the mid-5th century (c. 450 AD) after the end of the Roman rule (410 AD), during which invading continental Germanic settlers “destroyed the fabric of Roman British society in a colonization movement” (Schrijver 2014: 15). After the Anglo-Saxons settled on the island various dialects of Old English emerged, which roughly corresponded to the continuation of the dialects spoken by the different invading Germanic groups, that is the Angles, the Saxons, and the Jutes. The English language itself is first attested in the second half of the 7th century. The traditionally accepted view is that four main dialects of Old English existed, which are: West Saxon, Mercian³³, Northumbrian, and Kentish (cf. Hogg 2011: 4), of which Northumbrian and

³³ Thomason & Kaufman (1991: 265-266) further divide Mercian into West and North Mercian dialects, of which North Mercian is attested only in one gloss, and assert that the Middle English dialects could not have

Mercian are collectively known as the Anglian dialects (after the Angles) and Kentish is typically associated with the Jute settlers. Historical linguistic data suggests that even at the time of the Germanic conquest, there might have been “dialectal nonuniformities” in English (Thomason & Kaufman 1991: 265), and Old English dialects can also be problematic to describe and delineate due to the nature of their historical sources. The main problem with the description of OE dialects arises from the scarcity of available written documents and that their descriptions are based on the shared characteristics of certain manuscripts (Colman 2012: 74, cf. also Hogg 2011: 5-7).

The OE dialect situation cannot be considered to be one in which a well-definable standard variety existed, or that the dialects themselves were free of any variation. Languages at a pre-nationalistic and ‘pre-printing’ stage of their history, such as in the case of OE, cannot be considered standardized, monolithic entities³⁴. It should also be noted though that it would be equally wrong to assume that languages in their present-day form constitute invariant units without any regional, social and individual variation. During their historical development, various English dialects acted as what could be called ‘standard’ from a modern perspective, which meant that as power shifted from one kingdom to the other so did the perceived prestige of the given dialect adjust accordingly.

The Anglo-Saxons started settling from the mid-5th century onwards in England where by the 7th century they had established the so-called Heptarchy, which were more like individual spheres of influence of rulers rather than actual kingdoms with clear borders, as aggressive rulers would often seize control of neighboring territories which often led to a fluctuation in borders and the size of dominions (Jones 1990: 48, Colman 2012: 75). Language-wise, in the 6th century first the Kentish dialect and the Kent area was the most influential one, due to the fact that St. Augustine’s mission first landed there in 597. With the recognition of the archbishopric of Canterbury and with Æðelberht I (c. 580 – c. 616) being the first Anglo-Saxon king to convert to Christianity and to be baptized, Kent gained prestige in the early era of Anglo-Saxon England (Cannon & Hargreaves 2001: 15, 18). Then, with the spread of Christianization, from the 7th century the Northumbrian variety was the dominant one. This is the area where York, Jarrow, and Lindisfarne are located, which were outstanding centers of culture and education, and were also the three most important sites of the early

developed solely from these OE varieties, and propose the existence of three more unattested Old English dialects: East Saxon, East Anglian, and East Mercian.

³⁴ It should also be kept in mind that neither can they be considered monolithic in the present day. It would be an essentialist, and ultimately wrong, standpoint to claim that after standardization and codification languages become devoid of any variation.

phases of the Viking conquest at the end of the 8th century. It is no coincidence either that the leading role of Northumbria and the Northumbrian dialect declined with the Scandinavian invasion, and shifted briefly to Mercia, then from the 9th century further south to Wessex and to the West Saxon dialect. The Modern English standard variety is also one that emerged on the basis of the variety that developed from West Saxon, as the three most important, culturally and administratively significant English settlements are in the south: London, Cambridge, and Oxford.

As it was discussed in previous subchapters, Old English served as the superstrate to a Celtic substrate and that it borrowed Latin lexical elements that entered the language either directly or mediated through Late British Celtic. Latin and Old English were used as the primary written languages, where Latin was the language of the church and education, and OE was the everyday language. With the end of Roman rule and with the Germanic conquest, “spoken Latin became extinct in Britain” (Schrijver 2014: 15-16). As opposed to its influence on British Celtic, Latin had a less significant effect on Old English due to the reason that speakers of OE were not under direct rule by Romans like the Celts were (cf. Schrijver 2002: 87). Latin for OE was a prestige language whose status derived from its use in the ecclesiastic and educational domains, and in Vennemann’s (2011) terminology, as discussed in Section 2.3., was most likely a prestrate in relation to Old English.

Besides functioning as a superstrate to British Celtic, Old English was the substrate of two invading languages: Old Norse (between the late 8th century and the mid-11th century) and Norman French (between the mid-11th century and the 13th century), both of which left a significant mark on English. The characterization of Old Norse as a superstrate to Old English has generated much discussion in the field of contact linguistics, as it has been discussed throughout this dissertation, because ON can also be considered to have been an adstrate of OE, due to the widespread settlement of Scandinavians and the vast amount of borrowed Norse lexical items in OE, as well as the impact of Old Norse on OE toponyms. Without doubt, however, Norman French served as a superstrate to English, which is evidenced by the thousands of loanwords that entered English and also by the negligible extent of influence that Norman French had on English place-names.

In conclusion, Old English evolved from the language brought to the island by the Germanic invaders, and it absorbed structural features from the language of the subordinated Celtic population and borrowed heavily from Old Norse and the superordinate Norman

French. Due to the eventful history of the language, multilingualism was probably also widespread in Anglo-Saxon England³⁵.

4.5. The Viking invasion of England

The fact that the main focus of this dissertation is the hybrid place-names of England with special attention to English-Scandinavian hybrid formations would warrant the inclusion of a separate chapter dedicated to the socio-historical, linguistic, and sociolinguistic background of the contact situation in which these names emerged. However, due to the constraints of space, the English-Scandinavian contact situation will only be dealt with in a subchapter that is somewhat longer than the others. It will be argued here that the relationship between Old English and Old Norse was more of an adstratal nature, that the Scandinavians presumably settled in England in two waves and in considerable numbers, and that mutual intelligibility and some form of bilingualism must have existed between the English and the Scandinavian population.

Providing a thorough analysis of the background and nature of the Norse expansion in England and Europe would lie far outside the scope and purpose of this dissertation, but since the Viking Age features as a prominent theme, and language contact phenomena cannot be analyzed properly without exploring the relevant historical context, I will present a summary of the most important pieces of information available on this matter. The discussion will not be, and is not meant to be, comprehensive.

In the late 8th century, the Anglo-Saxons, now settled in the British Isles, were subject to raids by the Norsemen from Scandinavia, but these were initially hardly ever more than just sporadic attacks in the northern coastal areas. It is a generally accepted fact that Scandinavian contact with the English began in 793³⁶, when raiding Vikings sacked the monastery of Lindisfarne, which also marked the beginning of the Viking Age in the whole of Europe. Seventy years later this attack was followed by two distinct waves of raids, some 150 years apart starting with the campaigns of the *micel here*, the great army of the Danes in the autumn of 865. This raid, unlike the earlier sacking of the monastery, was accompanied by settlement, especially after the battle of Edington, when Alfred, king of Wessex made peace with the

³⁵ However, Cain (2016: 81) raises awareness to the issue that medieval multilingualism is almost uniformly associated with the presence of linguistic borrowing which is only indirect evidence. In reality, medieval multilingualism means an “awareness of and use of multiple languages in a variety of texts” (ibid.) but actual descriptions of multilingual practices in OE times are found very sparsely in manuscripts.

³⁶ Norwegian Vikings made very brief contact with the English before this in 787 when three of their boats arrived in Dorset (Loyn 1994: 38-39, Moskowich 2012: 21-23).

Danes in 878 with the conclusion of the Treaty of Wedmore and the establishment of the Danelaw.

Prior to this peace treaty, from the mid-9th century onwards “[m]uch more widespread plundering by large armies marked the [...] period, and this resulted in extensive settlements and the establishment [...] of Norse institutions” (Dawson 2003: 2). By the late 9th century the Danes had taken over Northumbria, where many of the raiders settled. Furthermore, the Norsemen attacked Mercia, East Anglia, London (which they managed to take over) and Wessex. Norse raids, however, still did not cease, yet the English were slowly regaining their territories and by the mid-10th century almost all of England was under English rule (Jones 1990: 60). Then, in 991 a new wave of Viking raids began from the southern coast of Wessex and by 1016 the Danes had control over the whole of England, now ruled by Canute of Denmark, who reigned until 1035. The Anglo-Norse state finally ended with William of Normandy’s conquest in 1066 (Robinson 1992: 141).

Moskowich (2012: 15-39) dismisses those theories which claim that the motivation behind the Viking eruption was solely pillaging and raiding, and proposes that the Viking expansion happened as a result of a much more complex series of historical and socioeconomic events that include the search for more and better quality arable land, which was of rather short supply in the harsh conditions of Scandinavia, the superiority of Norse seafaring and the disappearance of the Danish royal dynasty in 854. Moskowich also discusses the second wave of Viking migration that took place shortly after the first group of Norse warriors, some one thousand men, who were defeated in England, began to settle down and assimilate. This second wave of migration was of a much larger scale and also included peasants and women, and, as Moskowich suggests, presumably it had a more pronounced linguistic influence than the first wave. This chain of events seems to be generally accepted by the prevailing scholarly position

4.5.1. The extent and the nature of the Viking settlement

As it will become clear from the following discussion, the scholarly literature seems to be somewhat divided over the exact nature and extent of the Viking settlement in England. The division runs along the question of whether a large scale migration and settlement took place or whether it was only a small group of soldiers of the Great Danish Army (the *micel here*) who settled in England. Arguments abound on either side of the issue, but the size and

makeup of the place-name material is often cited as a decisive circumstance in favor of a massive migration.

Sawyer in his very influential monograph on Viking Age England (1971) argues in favor of the second configuration, namely that only soldiers settled in England. The counterargument, however, proves to be more convincing: practically the entirety of Fellows-Jensen's scholarly oeuvre (but especially 1972, 1978a, 1978b, 1980, 1985), Loyn (1994), Cameron (1996), Townend (2002), all seem to be in agreement that Scandinavian settlement in the Danelaw was the result of extensive migration, and from the place-name data of the present dissertation we can also draw similar conclusions. Wherever the Scandinavians roamed during their explorations they assimilated into the local population, for instance in Normandy, Ireland, etc., and did not retain their language (Thomason & Kaufman 1991: 267), but they remained a very strong component of Northern England's population. Thomason & Kaufman (*ibid.*) assert that Scandinavians settled in North and East England between 865 and 955 and that their language survived for no more than two generations after 955.

It has been suggested in the literature that place-names with Norse generics and Scandinavian personal names as their specifics indicate new settlements of Norse invaders, and were coined first. Hybrid names containing an ON personal name and an English generic are hypothesized to have been taken over at an early stage by Scandinavians and partially renamed. In general, names in *bý* 'village' tend to occupy the best available uncultivated land, followed by names in *þorp* 'outlying farmstead' that do not necessarily indicate settlement by Scandinavians but are only indicative of Scandinavian influence (Jones 1990: 60-61). It has also been argued (Jones *ibid.*) that appellative names in *bý* represent English settlements renamed by Norse invaders on account of them occupying similar sites as Anglo-Saxon settlements, leading to the conclusion that they arose as a result of the fragmentation of Anglo-Saxon estates. Finally, names in *þveit* 'clearing' are taken to be representative of settlements established by Scandinavians on virgin land. – Fellows-Jensen?

Ringe & Eska (2013: 74) reason that Scandinavian, or Scandinavian-influenced, place-names in the northern territories of England could not have been produced by an English-speaking population by simply renaming the already existing settlements under pressure from the Norse population³⁷. Cox (1980: 48-49) also adduces further pieces of evidence against the theory of renamed settlements, mainly recapitulating Cameron's (1975) observations. Using a

³⁷ The authors also bring a very persuasive argument in favor of this view, namely that if Scandinavian toponyms in England came about through English speakers giving new names to their settlements, then Norman French should have also had such a powerful impact on English toponyms, which it did not.

combination of examining the place-name material, the quality of soil, and the topography of the area, his conclusion is that apart from pure English name formations, the best sites are occupied by settlements that have a hybrid place-name containing a Scandinavian personal name and Old English generic. These settlements are likely to be ones which changed ownership in the aftermath of the Scandinavian conquest (cf. also Fellows-Jensen 1972: xx).

It is indeed very unlikely that such a large number of villages all across the areas under Scandinavian rule would be renamed by English speakers, therefore we can assume that they are either new coinages or they arose in a (semi-)bilingual situation, and that they were created by the Scandinavian speakers. Furthermore, based on the place-name evidence, Ringe & Eska (*ibid.*) also conclude that “significant numbers of speakers of Old Norse settled in England following the Scandinavian invasions.” This standpoint, contrary to Sawyer’s (1971) postulation of a relatively small initial settlement, seems to be the generally accepted view in the literature, and can be corroborated by the wide variety of Old Norse generics employed in pure Scandinavian formations and English-Scandinavian hybrid names (for details see the discussion in Section 6.7 below). A dual wave of influx would also point in the direction that the Norse army came to conquer and settle and not solely to pillage.

It is postulated here that the massive settlement of free peasants brought about a state of affairs in which neither the speakers of English, nor those of Old Norse were in a subordinate or superordinate position. In the Danelaw area Danish law prevailed (cf. Jones 1990: 58), yet there were no clearly observed boundaries of identity, ethnicity and nationality between the English and the Danes (Hadley 2002: 52). Further pieces of evidence can be provided by the sheer amount of Norse-related place-names (be they purely Scandinavian, Scandinavianized or English-Scandinavian mixed ones), which would hint at a larger-scale settlement and population mixing, disregarding whether or not this was the result of internal colonization, as discussed above. This situation is in stark contrast with the one that was the result of the Norman Conquest, which meant a complete replacement of the Anglo-Saxon ruling class with an aristocracy of Norman French origin, and in this case Norman French, the language of the invaders was clearly and purely superstratal.

Finally, a word of caution should be in order here, however. Overly relying on toponymic data for analyzing population movement and drawing far-reaching conclusions about the nature of settlement of a given population group can have serious pitfalls and caveats. As Townend (2000: 96) observes “[i]t is widely recognized that the number of Norse place-names and loanwords cannot be directly correlated with the number of Scandinavian settlers, but only with the extent of Norse linguistic influence” meaning that toponyms are

only indicative of the fact that the contact situation was intimate with far-reaching consequences because Scandinavian naming patterns could have exerted influence on English place-names in areas without extensive Norse settlements. It is also important to bear in mind, as Muir (1998: 1-3) points out, that if a settlement bears a certain name it does not necessarily mean that the given settlement was established or populated by a population whose language contains the place-name elements found in that toponym.

4.5.2. Language in Viking Britain

The type of the contact of English with Scandinavian and the influence of the latter on the former was intimate, pervasive and most probably dominantly superstratal with adstrate-like characteristics, and based on the linguistic evidence, Townend (2000: 96) also concludes that “Old English and Old Norse were adstratal in Viking Age England, [...] they enjoyed roughly equal prestige”. Naturally the extent and intensity of the English-Scandinavian contact situation varied from region to region as well as across time, and we cannot speak of a distinct, solid wave of influence which uniformly affected the English language,³⁸ yet the overall superstratal influence is clearly indicated by the fact that the trigger for the emergence of the contact situation was the conquest by the Vikings and the language of the victorious group is typically a superstrate language (Thomason & Kaufman 1991: 116, Hickey 2010a: 7, Vennemann 2011: 218). Furthermore, this effect is also observable in the breadth and variety of Scandinavian loanwords that entered the English language, which on occasion replaced native English terms, other times introduced new words to fill lexical gaps and in some cases led to the emergence of etymological doublets. This rather vast amount of borrowed lexical items, as demonstrated by Moskowich (2012: 110-122), belongs to quite a wide variety of semantic fields, therefore they are not exclusively or dominantly found in the rather broad and elusive category of “everyday life”, as it is very often posited.

Vennemann (2011: 242) and Lutz (2013: 567) also raise awareness about the issue that many of the Scandinavian loans in the field of law and administration were superseded by words borrowed from French following the Norman Conquest, which obscured the evidence of the Norse domination over the political system (at least in the Danelaw), thus one would be

³⁸ In this situation, one also has to bear in mind that neither the Old Norse nor the Old English language constituted a standardized, uniform entity, as there existed regional variation and dialectal differences in both cases. Furthermore, the Scandinavian loanwords found in Modern English are the ones which have survived the Norman French invasion and its massive linguistic impact (Lutz 2012) and also passed standardization after becoming part of the standard variety, largely based on the West Saxon dialect and Southern English.

led to erroneously believe that the language of the Scandinavians did not exert such a significant superstratal effect. Similarly, Werner (1991: 381) also observes that the Old English period is invisible in terms of “language-mixing” with Scandinavian due to the fact that in assessing the Scandinavian influence we can only rely on what percolated down to the Middle English period. In keeping with Vennemann’s and Lutz’s point above, many unrecorded Norse loanwords must have vanished under Norman rule. Evidence for the most likely, partially adstrate-like nature of the Anglo-Scandinavian contact situation can be found if one subscribes to the (quite probable) scenario of a secondary, peasant migration, which followed the campaigns and conquests of the *micel here* (cf. Moskowich 2012: 15-39, as referenced and discussed in 5.1 above).

Werner (1991: 379-382) raises awareness that in accounts of the contact between English and Scandinavian, the umbrella term ‘Old English’ is misleadingly used for referring to the language of England, and similarly ‘Old Norse’ is misleadingly used for denoting the language of the Scandinavians. The reality is that later Mercian and Northumbrian OE was in contact with Proto (or Runic) Nordic, therefore the situation should be labelled as the influence of “Proto-Norse on Northern Old English.” The problem arises from the fact that there are no extant textual records of Viking Age northern OE (cf. also Ringe & Eska 2013: 74) and no records of “Anglo-Danish” texts either, and Scandinavian elements only appear in Middle English in the 12th century, by which time the contact had long been over.

The Norse rule, or Norse presence, lasting for at least two hundred years left its mark on the language as well. Mostly the lexicon and the phonological system of the language was influenced in this period, since a great number of borrowings and loan translations entered the English vocabulary at this time and certain phonological features were also introduced to English³⁹. There are instances of borrowing a word from Norse for which there is no English counterpart as well as replacing an existing English word with a loanword, such as *take* < *taka* ‘to take’ which replaced Old English *niman* (Dawson 2003: 4).

The signs of Norse influenced language change have already begun showing up in the Late Old English era, around the 10th century, chiefly in the Northern dialects. However, their exact appearance is difficult to track down and pinpoint, due to the fact that they were restricted to the Northumbrian and Mercian varieties and thus were regarded as dialectisms and could not enter the written forms of the West Saxon variety, which came closest to what

³⁹ This phonological influence, however, was different than the one in Roman Britain where the Brittonic language incorporated phonological changes induced by Latin. That was the result of deliberate Romanization, whereas ON influence is due to the adstratal relationship of ON and OE.

could be considered a ‘standard’ and which by about the year 900 came to be the most widely used written variety (Thomason & Kaufman 1991: 266). However, these features slowly did manage to find their ways into the southern dialects and with the Norman conquest of 1066 when all “regional varieties [...] were leveled to a low but equal sociolinguistic prestige” these northern features could enter even the language of written London English (Rot 1992: 186-187).

Besides the British Isles, the Vikings settled during their explorations in Russia and Normandy as well (the name Normandy itself derives from the name of the Norsemen), and although in these latter two regions they were linguistically quickly assimilated to the population, they were able to make a lasting impact on English (Adams & Mallory 1997: 219), affecting its phonology, morphology and lexicon. The contact between Old English and Old Norse is characterized as intimate borrowing, because several levels of the language were affected and the loanwords are not restricted to content words only, but pronouns were also borrowed and phonology and pronunciation was also influenced. Bilingualism is also likely to have existed between ON and OE on the basis of hybrid place-names in which loan-translated elements can be found and also on the basis of phonological influence which can be taken to be an indicator of bilingualism (Thomason & Kaufman 1991: 37).

In phonology, for instance, the Proto-Germanic *sk cluster yielded sc /ʃ/ in Old English (which in Modern English is represented by <sh>) in every position (although not entirely without exceptions), but in Old Norse it was retained as /sk/. When Old Norse got in contact with Old English, the sk cluster was reintroduced through borrowing, and sometimes it gave rise to etymological doublets, i.e. pairs of words which stem from the same source with one word inherited and the other one borrowed. Such a doublet is the pair *shirt* and *skirt*, both of which go back to PGmc. *skurtjon, which developed into Old English *scyrte* and Old Norse *skyrta*.

Another way in which Scandinavian loans influenced English phonology was the pronunciation of *k* and *g*. In Old English *k* was pronounced as /kʰ/ (i.e. it was palatalized) if it stood before or after front vowels *i* or *e* and *g* was palatalized to /j/ in the same environment. In Old Norse, on the other hand, there was no such palatalization (though it did exist but under different circumstances and with different outcomes than in Old English). The influence of Old Norse in this case can be seen in the pronunciation of words such as *kid* or *give* both of which are pronounced without palatalization, however in Old English the form *giefan* was realized as /jievən/ which is reflected in the Middle English spelling *yiven*. This word in Old Norse appears as *gefa* and is realized as /geva/ and the spelling and pronunciation of Modern

English *give* is attributed to Scandinavian influence during the Middle English period. These phonological changes are important for the analysis of place-names, because a significant number of English toponyms, the so-called Scandinavianized names, exhibit Norse phonological features such as the ones discussed above.

Moving on from phonology, we can find signs of Norse influence in the lexicon as well, both among content words and function words. Two groups can be differentiated within content words:

- (i) borrowing new words which are not found in Old English and
- (ii) borrowing words which will replace Old English forms.

Examples for the former case (i.e. necessity borrowings) include some of the most frequent words of Modern English, e.g.

1. *fellow* < ON *felagi*, a darkened compound word of *fe* ‘money’ and a verbal base meaning ‘to lay’, originally used in the sense “one who puts down money with another in a joint venture”.

Examples for the latter case (i.e. luxury borrowings) include

1. *window* < ON *vindauga* (a darkened compound of *vindr* ‘wind’ and *auga* ‘eye’) which replaced OE *eagbyrl*, itself also a compound: ‘eye-hole’ (the second element *-byrl* can be found in for example *nostril*) or
2. *sky* < ON *ský* ‘cloud’ which replaced OE *heofon* ‘heaven, sky’ (note that in modern Germanic languages cognates of OE *heofon* are still used for denoting both ‘heaven’ and ‘sky’, e.g. German *Himmel*).

In conclusion, the Old Norse language had a far-reaching impact on Old English which is evidenced by the numerous loanwords (which are often luxury borrowings, including the famous example of the 3rd person plural personal pronouns *they*, *them*, and *their*) and the phonological changes and alterations that are the result of this contact situation and are also observable in toponyms. ON and OE were structurally very similar and a degree of mutual intelligibility very likely existed. Assessing the exact nature of the stratal relationship though is still problematic and subject to debate.

4.6. The end of the Viking era and the Norman Conquest

The end of the entire Viking period in England came with the Battle of Hasting and the subsequent Norman Conquest of 1066. Following the battle, between 1066 and 1070 Normans with Breton and Flemish allies conquered England (Thomason & Kaufman 1991:

267) and William I (or William the Conqueror, and William the Bastard, especially in the Norse tradition, 1066 – 1087) ascended the throne. In the aftermath of the Norman Conquest of 1066, most of the English aristocracy and clergy was replaced by men primarily of Norman French origin. Given that they occupied the higher classes of society they numbered considerably fewer than the English-speaking population. Thomason & Kaufman (1991: 268) estimate the number of Norman fief-holders to be around 20,000 while the population of England was around 1.5-2 million, while Fellows-Jensen (1980: 196) argues that “the total immigration is unlikely to have exceeded 10,000” and the “number of foreigners who were granted land by William was less than 2,000”, so it could be assumed that the total population of Norman French speakers, or rather new settlers in England was somewhere between 10,000 and 20,000, perhaps tending more towards the lower figure (cf. Clanchy 2014: 45).

The result of the conquest was not exhausted by the overhaul of the English ruling class, but it also brought about changes in the English language, mostly in its lexical system. After the Norman Conquest, the English literary language vanished for a century under Norman rule, although English continued to be spoken, but it only resurfaced as a written language in the mid-12th century. Thomason & Kaufman (1991: 268) reason that prior to the loss of Normandy in 1204 under the reign of John Lackland⁴⁰, the incumbent English king was also the Duke of Normandy, and the high ranking nobility split their time between England and France. Therefore, they had no incentive to learn English or require the local population to learn French, but this changed after 1204 when the aristocracy started using English more extensively, thereby giving a boost to the English language. A large number of Norman French loanwords entered English during this dormancy period, and the Scandinavian elements also appear in large numbers only in Middle English, even though the contact between OE(/ME) and ON had long ceased by then. By the time Norse loanwords surfaced in Middle English, they had already been integrated into the language (cf. Werner 1991: 380-381, Moskowich 2012: xx).

The contact situation between Norman French and English is the typical example of superstratal language shift, involving heavy borrowing and lexical importation, with limited structural influence and negligible effects on place-names. English received the heaviest influence from NF in the lexical part of the language, but certain phonetic features of ModE are also the result of contact with French. Perhaps the most well-known example of phonological interference would be the phonemization of [v], [z], and [ð] which in OE only

⁴⁰ 1199 – 1216 (Cannon & Hargreaves 2001: 196)

existed as allophonic variants of voiceless /f/, /s/, and /θ/. In this context, Norman French functioned as the superstrate because it was the language of the conquerors and thus had higher prestige, and Old and Middle English functioned as the substrate. Bilingualism can also be posited to have existed between Norman French and Middle English, which probably emerged relatively soon after the Norman takeover, as Clark (1992b: 121) also asserts that Norman French monolingualism could not have survived for more than three generations after the Conquest.

William the Conqueror ordered the Domesday Book to be compiled, for which the survey was made in 1086, and its main purpose was to have a record of taxable estates. The Domesday Book was an ambitious endeavor involving multilingual interaction, as it was written in Latin for a French-speaking elite about England and its mostly English-speaking residents, with information elicited by French-speaking lords (Baxter 2011: 271-273) and compiled by a primary scribe who was “English-trained and probably English-speaking” (Clark 1992b: 122-123, cf. also Roffe 2000: 73 who asserts that the scribe was northern and at least bilingual if not native English, and Baxter 2011: 292). The DB itself consists of three main sections comprising returns of the inquest: (i) the Exon Domesday, covering the south-west shires of Cornwall, Devon, Somerset, Dorset, and Wiltshire, which was incorporated into the main volume (Great Domesday), (ii) the Little Domesday, which was never incorporated into the Great Domesday covering Essex, Norfolk, and Suffolk, and (iii) the Great Domesday, covering the rest of England with the exception of London and Winchester (Roffe 2000: 75, 89, Baxter 2011: 288-289).

The book is an invaluable resource for place-name research, as many toponyms are first attested in it. However, a drawback of the book is the often inconsistent spellings, which are the result of the multilingual context in which the DB came to life. As it was discussed in the previous chapter, the influence of an “Anglo-Norman” scribe was for long held to have been the reason for irregular changes in toponyms, but later research disproved this theory. The Great Domesday Book mostly rendered OE names in a consistent manner, and it should be noted that the Little Domesday Book’s orthography is closer to the original Old English spellings than the Great Domesday (Roffe 2000: 89). The DB’s importance lies in the fact that it was compiled by scribes who had a good command of the English language and treated names appropriately, therefore it is a reliable historical source of toponyms. The problem that is still present though is that in many cases no records of pre-DB forms survive which would be necessary for establishing how exactly the hybrid names came about.

In conclusion, the influence of Norman French on English can mostly be measured in the high number of loanwords that it supplied. Given that NF was a language of high prestige, due to the political supremacy of the French ruling class, thousands of new lexical items entered Middle English by the 15th century. In accordance with the heavy lexical importation, the language shift that occurred in this case was clearly superstratal. Due to the dormant period of OE soon after the Conquest, Old Norse loanwords also first surface only under Norman rule, even though they had entered the language during the Old English stage. After the loss of Normandy in the early 13th century, Middle English is reinstated as a written language after which French elements still kept on to be incorporated by English.

Table 2 on the following page summarizes the stratal relationships between various language pairs throughout British history along with the indication of the nature of influence they exerted on each other. The types of place-name formations characteristic of each of the layers discussed here will be elaborated on in the next chapter.

	period	stratal relationship				shift	influence		
		sub-	super-	ad-	pre-		str	lex	pn
Pre-Celtic – Celtic	uncertain	Pre-Celt.	Celtic	/	/	SUB → SUPER	0	0	1
Celtic – Latin LZ	43 AD – 410 AD	Celtic	Latin	/	/	SUB → SUPER	2	2	0
Celtic – Latin HZ	43 AD – 410 AD	Latin	Celtic	/	/	SUB → SUPER	2	2	0
English – Latin	450 AD –	/	/	/	Latin	NO SHIFT	1	2	0
English – Celtic	450 AD –	Celtic	English	/	/	SUB → SUPER	2	0	0
English – Norse	793 – 1066	English*	Norse*	English, Norse	/	SUPER → SUB	2	3	3
English – N. French	1066 – 13 th c.	English	French	/	/	SUPER → SUB	1	3	0

Table 2. Summary of the stratal relationship of languages in Britain and their influence on each other

Legend:

str: structural influence, *lex*: lexical influence, *pn*: place-name influence, *LZ*: Lowland Zone, *HZ*: Highland Zone

0: no or negligible influence, *1*: light to moderate influence, *2*: moderate to heavy influence, *3*: heavy influence

The definitions of strata are to be understood as per Vennemann (2011), and Tristram (2007: 198) for adstrate. The term ‘English’ is to be understood as a general cover term for any diachronic variety of English, as the table is meant to be the schematic representation of an aggregate of influences. Influence refers to the influence exerted by the shifting language or if no shifting occurred than by the non-English language, with the exception of LZ where the structural influence was phonological and morphosyntactic from Vulgar Latin on British Celtic.

V. Stratification of British Place-names and Hybrid Names

This chapter, which comprises the dissertation's empirical background and the bulk of the results of the empirical analysis, presents the corpus that serves as the basis of the investigation and also the data that it can provide for English toponyms of various origins (i.e. Celtic, Latin, English, Norse, and French). These place-name formations are surveyed in the chronological order outlined above. Presented against the historical background and context in which the names were created, the chapter surveys the occurrence and usage patterns, and origins of the specifics and generics of each linguistic layer. The organizational principles of the corpus are outlined and the rationale behind the tagging method is provided.

5.1. The corpus of settlement names

The place-name data for the analysis was chiefly drawn from the place-name dictionaries of Ekwall (1980) and Mills (1998, 2011), with additional data of Scandinavianized forms, lost villages, deserted and depopulated settlements being drawn from Fellows-Jensen's three major regional studies of Scandinavian settlement names in Yorkshire (1972), the East Midlands (1978a), and the North West (1985). The area of England was surveyed only, names from Cornwall, Wales, Scotland and Ireland do not form part of the corpus. Altogether 10,311 relevant place-names have been collected, from which a special corpus has been compiled for the purposes of the present study. Due to the reasons described in Section 1.5., the name-material gathered in the corpus is a smaller, but still sizeable, subset of the entire historically possible set of English toponyms. Despite this shortcoming, the corpus is a self-contained, representative dataset, and the conclusions drawn in the dissertation shall be understood in the light of these limitations, meaning that no far-reaching conclusions are possible to be drawn, even though the dataset can be taken to represent typical tendencies.

The entirety of the name-material was entered into an XML (Extensible Markup Language) corpus, with the following 12 types of tags applied to each entry:

<code><fullname></code>	present-day form of the place-name
<code><att.formN></code>	historically attested form, with an identifying number added for each date of attestation
<code><meaning></code>	original meaning of the place-name

<spec>	the specific of the place-name
<speclg>	the language of the specific
<gen>	the generic of the place-name
<genlg>	the language of the generic
<spectype>	type of the specific (appellative, personal name, river, etc.)
<sourceN>	source/date of attestation, with a number corresponding to <att.formN>
<region>	geographic region in which the place-name is found
<note>	special notes for the given place-name (e.g. uncertainty of origin, replacement of certain elements, etc.)
<scand>	tag used for indicating if a given place-name underwent Scandinavianization

Depending on the number of attested forms recorded in the corpus, and whether or not a name was Scandinavianized or has some special notes attached, each name consists of 10 to 20 lines of data, which means that the entire corpus totals at around 125,000 lines of XML data. The first approximately 10% of the name-material (altogether 1115 names) was entered into the corpus manually, which served a dual purpose. On the one hand, it was done this way so that the potential problems and shortcomings of the tagging system, the markup language, and the transformational commands could come to light early on in the process of corpus building and they can be remedied before the corpus reaches its full extent. This way the tagging principles could also be fine-tuned without having to modify a huge body of toponyms. In brief, this initial 10% served as a ‘sandbox corpus’ (with entries ranging from *Abberley* (WORCS) to *Billesley* (WARWICKS)) in which all the minute details of tagging and transformation could be ironed out. On the other hand, this method was employed as a means of economizing on resources, because the remaining 90% was fed into the corpus in a semi-automatized way based on the experimental smaller corpus. Due to the size of the corpus and the sometimes inconsistent formatting in the source material, the resulting dataset was checked for any mistakes and misplaced tags with the search and replace functions of Excel 2016 and Notepad++ and were corrected if needed. The e-book version of Mills (1998, 2011) was cleaned up, stripped of all entries deemed unnecessary or unfit for the present

investigation⁴¹, and transplanted into an Excel worksheet, where each column represented a separate tag of the place-name records in the XML file. At this stage any special characters in the input text (such as macrons indicating vowel length) were replaced by regular vowel letters with accents so that they can be processed by the software. This worksheet was then merged with the one used for the paper that can be considered the embryonic version of this dissertation (for that analysis see Fekete 2015) and which contained the entries on depopulated settlements and lost settlement names from SSNY, SSNEM, and SSNNW. The Excel file was then exported as an XML format file and it was merged with the ‘sandbox’ pilot corpus, and so the whole body of English place-names emerged. Figure 3 below shows a snippet of the first two entries of the corpus in an XML format.

```

1  <?xml version="1.0" encoding="utf-8"?>
2  <placenames>
3  <pn>
4      <fullname>Abberley</fullname>
5      <att.form>Edboldelege</att.form>
6      <meaning>Woodland clearing of a man called Eadbeald</meaning>
7      <spec>Eadbeald</spec>
8      <speclg>OE</speclg>
9      <gen>léah</gen>
10     <genlg>OE</genlg>
11     <spectype>pers</spectype>
12     <source>DB</source>
13     <ref>MB52</ref>
14     <region>Worcs</region>
15     <note />
16     <scand />
17 </pn>
18 <pn>
19     <fullname>Abberton</fullname>
20     <att.form>Edburgetuna</att.form>
21     <meaning>Farmstead or estate of a woman called Eadburh</meaning>
22     <spec>Eadburh</spec>
23     <speclg>OE</speclg>
24     <gen>tún</gen>
25     <genlg>OE</genlg>
26     <spectype>pers</spectype>
27     <source>DB</source>
28     <ref>MB52</ref>
29     <region>Essex</region>
30     <note />
31     <scand />
32 </pn>

```

Figure 3. The first two entries of the place-name corpus in Notepad++

⁴¹ These were purely Gaelic and Brittonic place-names, and they were excluded from the corpus because (i) the main focus of the dissertation is the survey of toponyms occurring in England, and these names occur primarily in Ireland, Scotland, and Wales, and (ii) Mills’ dictionaries (1998, 2011) do not provide adequate descriptions and etymological background information for them unusable for this analysis. Even though it would be desirable to include Gaelic place-names in such a study for reasons of comparability and in the interest of a fuller treatment of the material, doing so would broaden the dissertation’s scope too much. Those Celtic/Gaelic names, however, which are relevant for its current scope are included in the analysis.

The use of XML for the compilation of the corpus was chosen because it is a highly flexible and lightweight construction which is both machine-readable and human-readable, completely customizable and allows user-defined tags to be applied. A further advantage of XML is that the tags do not have to follow a strict order, which means that certain tags can be left empty or can be omitted entirely, and the sequence in which the tags are entered into the file is irrelevant, because the transformation script will be able to find the value of any tag that is specified regardless of where the tag is located within the file.

5.2. Methods of analysis

The corpus was analyzed from a chiefly quantitative perspective, focusing on the distributions and frequencies of various types of place-name formations. The XML data was processed with Extensible Stylesheet Language Transformations (XSLT), which allows the user to extract data from the XML file according to user-defined criteria and in a user-defined format. The XML file was handled with Microsoft XML Notepad 2007, Notepad++ version 7.4.2., and TextPad version 8.1.1., all of which are freely available pieces of software⁴², while the XSL transformation style sheet was created with TextPad, the transformation was carried out in XML Notepad 2007, and the output of XSLT was transferred to and analyzed statistically in Microsoft Excel 2016. A number of different XSL stylesheets were prepared for the analysis. The main stylesheet contains all the place-names that occur in the corpus, separated into specifics and generics, with the language of each element indicated, as well as the values pertaining to most of the XML tags described above. A second stylesheet contains the frequencies of generics along with the language of those generics, while a third one contains the frequencies of specifics and their languages. A fourth stylesheet is used for investigating the areal spread of certain generics and specifics. In all cases, the place-names are primarily categorized by the language of their specifics.

The empirical analysis of the dissertation is grounded in the discipline of corpus linguistics. Corpus linguistics is a very useful and potent tool for analyzing language data (for an overview of the field see Szirmai 2005), but it should be kept in mind that while corpora can reveal trends and tendencies, corpora alone are not omniscient. They show trends of actual language use, and show what speakers *have* produced, but they do not show what else

⁴² Microsoft XML Notepad 2007 is available at the following link: <https://www.microsoft.com/en-us/download/details.aspx?id=7973> Notepad++ is available at the following website: <https://notepad-plus-plus.org/> and TextPad is available at <https://www.textpad.com>

could have been produced that would have been equally acceptable. Unfortunately, historical linguistics and historical code-switching has little else but corpora to rely on. The problem is that those texts that can be fed into corpora represent only a fraction of what could have been produced or what actually had been produced but was never written down or was simply lost over time.

The goal of the corpus analysis is to obtain large amounts of numerical data, which can be contrasted with each other and correlated with factors such as geographical location, place-name type (i.e. personal name or appellational specifics), and the stratal relationship of various languages discussed in the previous chapter. The frequency and percentage of occurrence, the type/token ratio, and the correlational patterns are the most important data points that were calculated and analyzed in the dissertation. All of the mathematical calculations were done in Microsoft Excel 2016 utilizing the standard, built-in functions of the program, and the tables that the XSL transformations generated were also imported into and analyzed in Excel 2016. The maps were created by the author in Microsoft Word 2016 based on a freely available template.

Finally, in some cases there might be discrepancies between the frequencies indicated in this analysis than those given in the preliminary study (Fekete 2015). The reason behind this is that in the previous study I relied on the earlier edition of Miller's dictionary (Miller 1998) while now I based my corpus on the more recent edition (Miller 2011), and, more importantly, the design of the corpus is different now. In Fekete (2015) I utilized Microsoft Excel to build and analyze my corpus, and now in the dissertation I am relying on XML and XSLT for the analysis which provides greater flexibility and precision for data extraction and other operations, therefore the results are more refined than they were in the previous iteration of the study.

5.3. A note on pre-Celtic names

Pre-Celtic names in Britain are usually considered to have been created by a substrate of speakers inhabiting the island before the arrival of the Celts, and are most frequently simplex toponyms to which later settlers often attached explanatory name elements, often resulting in tautological formations (Bölskei 2012: 151-152). The names of large topographical features, i.e. rivers and mountains in England are generally believed to be of such pre-Celtic origin (cf. Fellows-Jensen 1978b: 19, Cameron 1996: 36, Bölskei 2012: 155). Major European rivers are held to be “among the most ancient toponyms” and the pre-

Celtic river names were brought to Britain by early settlers to the island, as evidenced by the overarching similarities of European and British hydronymy (Hough 2012b: 213). These Old European hydronyms, and other pre-Celtic names, are generally accepted to be of Indo-European origin (cf. Kitson 1996, Mills 2011: xiv, Bölcskei 2012) but contrary to this assumption, some twenty years ago in an attempt to challenge this view, Vennemann (1994) put forward the new theory of the Vasconic substrate in Europe. According to his hypothesis, prior to the influx of Indo-Europeans, Europe was inhabited by speakers of Vasconic languages to which Basque is also believed to belong. The ancient, pre-Celtic river-names of Britain and Europe are considered to be the vestiges of this Old European, Vasconic substrate, which Vennemann assumes to be an agglutinating non-IE language.

The theory itself originates from Krahe, who uses the term ‘Old European’ to refer to the Western branch of Indo-European spoken before the time it split into the attested branches, including Celtic and Germanic. In essence, the core of the argumentation of the Old European Hydronymy theory, which is also accepted by Vennemann (1994: 218-221), is that there is a large number of recurrent suffixes across Europe and related hydronyms are spread out across a wide geographical area, which entails that there must be a connection among these names.

For the purposes of this dissertation, such a cursory mention of the theory is sufficient, as it must be mentioned in connection with any analysis of English toponyms. Since the analysis of pre-Celtic names is not at the focal point of the dissertation, only one important instance of such toponyms will be mentioned here. The name of *London* is assumed to be of pre-Celtic origin (Rivet & Smith 1982: 396-398, Mills 2011: 305) ‘place at the navigable or unfordable river’ (Mills 2011: 305), first attested around 115 AD in the form *Londinium*⁴³. Including the name *London*, the corpus lists altogether 24 instances that are labelled as “Celtic or pre-Celtic” of which only 8 are marked to be “pre-Celtic” without the possibility of them being of Celtic origin. Without exception all 24 such toponyms are formed from earlier river-names.

Besides the “pre-Celtic” label, place-names derived from river-names, hill-names, and tribal names are sometimes indicated to be ‘pre-English’ in the corpus, based on Mills (2011) which was its main source. This is quite an unfortunate and unclear term, as it could refer to either Celtic names or earlier pre-Celtic ones. They could have also been treated amongst the Celtic names, under the assumption that they were borrowed into Celtic and Celtic is the only

⁴³ For a discussion of the name’s origin and problems of providing a definitive etymology (a problem that is shared by most pre-Celtic names) see Rivet & Smith (1982: 396-398).

'pre-English' language that has a sizeable body of attested toponyms, but since they are not represented in high numbers, I will treat them briefly here.

Altogether 45 names are of "pre-English" origin in the corpus, none of which contains a personal name, and the overwhelming majority (41 names) derive from river-names. Out of the 41 toponyms that have a river-name as their specific, 7 are simplex names, meaning that the place-name's only element is a river-name specific, not other elements are added to it. The remaining 4 non-hydronymic names are 2 hill-names, 1 forest-name and 1 simplex toponym. If this figure is taken together with the 24 "pre-Celtic" names then it will give us 69 instances that could collectively be referred to as "pre-English."

In conclusion, the pre-Celtic names of Britain, or at least the river-names and names of large topographical features, can be assumed to be Indo-European in origin and to have formed a substrate, and those names that lack an adequate etymological explanation very likely originate from this substrate. Such toponyms and hydronyms survived because they were taken over by Celtic speakers and some were continued to be used by the Anglo-Saxons, even if in a slightly modified form. The figures above show that only a very small proportion of English place-names can be considered to be pre-Celtic in origin, however the ratio would probably be much higher in the case of river-names, hill-names and field-names. This indicates and confirms that these languages and their speakers were in a substratal position in relation to the Celtic population.

5.4. The Celtic layer

The vast majority of Celtic names found in England fall into the category of river-names or place-names derived from topographical features and river-names. Most of the names of rivers, hills and other topographical features in England are of Celtic or pre-Celtic origin (see also Cameron 1996: 35-38). The Celtic language in England following the Anglo-Saxon invasion functioned as a substratum, therefore Brittonic did not have a very significant influence on English place-names. This, however, does not mean that Celtic place-names are entirely absent in England, although it should be kept in mind that "nowhere in England is British [i.e. Celtic] influence on place-names paramount" (Clark 1992a: 480, cf. also Coates 2007). The only influence Celtic had on English toponyms is the river-names that were transferred to English place-names either in their simplex form (which occurred rarely) or accompanied by an additional OE generic.

Following Cameron (1996: 31-39) three main groups of Celtic names can be distinguished, in descending order of frequency⁴⁴: those that denote

- (i) rivers, mostly the large ones such as *Exe* (Devon) *Ouse* (Yorkshire), *Thames* (Gloucestershire), many of which could equally well have been created by a pre-Celtic population as the etymologies tend to be obscure and difficult to reconstruct definitively
- (ii) hills, such as *Bray* (Berkshire) ‘hill’, and *Crick* (Northamptonshire) from Celtic **creig* ‘rock, cliff’
- (iii) woods, such as *Chideock* (Dorset) < Celt. **cēdiōg* ‘wooded’, a derivative of **cēd* ‘wood’

Apart from these three main types, names for other geographical formations can also be found in low numbers. Celtic place-names are characterized by their reference to topographical features found in the vicinity of the settlements, which means that monolingual Celtic toponyms very often lack a categorizing generic that would be the equivalent of English *tūn* ‘village’ or Scandinavian *bý* ‘village’ and thus habitative names are very rare (Cameron 1996: 36). Instead, these settlements are named after rivers and hills which are transferred to denote the settlements themselves. As a result, Celtic toponyms and hydronyms in the overwhelming majority of cases refer to some topographical or hydrological features, such as ‘tall hill’ or ‘muddy water’ or simply mean ‘water’, ‘river’, ‘hill’, ‘mound’, etc., corresponding to the meaning of the river-name or hill-name. However, this does not mean that these place-names are always simplex ones, as a number of specifying elements can be added, which are very rarely of a habitative type.

The primary medium of British Celtic culture was oral, therefore the available data on pre-Germanic Celtic place-names is scant, and can mostly be found in foreign works on geography or history written in Greek or Latin, or conserved in Latin and Old English forms (Bölskei 2012: 156-157). Celtic contact with Germanic settlers had three logical outcomes concerning the fate of place-names: they were retained until the present day, lost in present-day English but retained in written sources, or incorporated into English names (Fellows-Jensen 1978b: 20). Incorporated names include for instance Latinized Celtic place-names that

⁴⁴ cf. Rivet & Smith (1982: 20-22) who distinguish six types of Celtic toponyms in descending order of frequency and age: (i) river-names, (ii) geographical features, including hills and woods, (iii) animal names, (iv) divine names, (v) ‘technical’ names, referring to agriculture and architecture, (vi) names related to human habitation. The first two categories here entirely overlap Cameron’s categories who divides hills and woods into two separate groups.

were extended in Old English, most frequently by the addition of the generic *ceaster*⁴⁵ ‘Roman camp’ or OE *burh* ‘fortification’ (often in the dative form *byrig*), but in the majority of cases they were mostly left untouched (Bölcskei 2012: 158). The higher number of surviving Celtic toponyms in certain areas (such as the North Riding of Yorkshire) reveals that the Germanic progress was slower there than in areas where fewer Celtic names have survived which allowed the Anglo-Saxon invaders to adopt and incorporate more Celtic names (Fellows-Jensen 1978b: 20, cf. Bölcskei 2012: 158).

As it was discussed in the previous chapter, in some areas bilingualism between Celtic and English must have existed, which allowed the substitution of Celtic place-name elements with synonymous English ones. For instance in the case of *Dover* (Kent) which in its original Celtic form was **dubrās* ‘the waters’, and the OE form is *Dofras* (*Dovere* in DB) which shows an original Celtic plural and the Old English form is also assumed to be a plural one (Cameron 1996: 32). I find this claim to be dubious as the forms are very similar to each other and similarity alone is not an indicator of how speakers analyze the given word, therefore the Celtic toponym could have been taken over as it was without it necessarily being analyzed as a plural form. No other source seems to corroborate the claim that the OE form of the name is to be seen as one that is marked for plurality (cf. Nicolaisen, Gelling & Richards 1970: 84, Ekwall 1980: 149, Mills 2011: 157). Still, the existence of bilingualism should not be ruled out, due to the reasons outlined in the previous chapter. It should be added, though, that it was not uniform across England and probably varied greatly with region and naturally with the extent and density of the Celtic population.

The occurrence of Celtic elements in the main areas of Anglo-Saxon settlement is very infrequent, mostly river-names and toponyms derived from river-names can be found here. Since this dissertation only deals with place-names, the Celtic river-names, hill-names and wood-names will be omitted from the analysis, however toponyms derived from such names will be included.

Concerning hybridization, English, Scandinavian and French elements can be found coupled with Celtic specifics (in this order of frequency), with the vast majority of generics coming from English. In some Celtic-English hybrid names such as *Bredon* (Worcestershire)

⁴⁵ The OE word *ceaster* itself is a borrowing of Latin *castra* ‘camp’ (singular *castrum*) which is treated by Mills (2011) and Cameron (1996) to be an English word and therefore count toponyms containing this generic amongst English names. Similarly, Celtic elements such as *cumb* ‘valley’ and *torr* ‘craggy hill’ are indicated to be English generics. Even though such categorizations disregard the etymological background, in keeping with the conventions of Mills (2011), I continued this practice in the compilation of my corpus, given that these elements were indeed borrowed and integrated into English, and are therefore the affected toponyms are not considered hybrid names in the strict sense.

< Celtic **bre* ‘hill’ + OE *dūn* ‘hill’ the Celtic element is complemented by a synonymous English generic, resulting in a tautological name. This raises an interesting question concerning hybridization and bilingualism. Typically those toponyms would be turned into tautological formations which are not intelligible to speakers of the reanalyzing language by adding a word that captures the most salient feature of the locality (cf. Eng. *Faroe Islands* < Faroese *Føroyar* ‘sheep islands’), and speakers who are aware of the meaning of the original name would avoid the addition of a superfluous element with a synonymous meaning. The other possible solution is that the original toponym had already been semantically obscured therefore the explanatory element was added.

In the corpus there are altogether 417 Celtic names (4% of the entire name material), of which 41 (9.3% of Celtic names) have a personal name specific, 97 (23.27%) have an appellative specific and the majority of names (265, which is 63.55% of all Celtic names) have a topographical feature in the function of the specific. Besides dithematic (and in some cases three-element) names, the corpus contains 14 simplex Celtic toponyms. Table 3 below summarizes the frequencies and distributions of Celtic names in the corpus that will be further analyzed in the following subchapters. The percentages given in the “Hybrid total” line are calculated from the “Grand total” line of the given type of specific.

<i>Celtic names</i> <i>N</i> = 417 (4%)		Personal name		Appellative		Topographical		Simplex	
	count	percent	count	percent	count	percent	count	percent	
<i>Monolingual</i>	2	0.48%	14	3.36%	136	32.61%	14	3.36%	
<i>Celtic-Latin</i>	1	0.24%	1	0.24%	0	0			
<i>Celtic-OE</i>	6	1.44%	70	16.79%	120	28.78%			
<i>Celtic-ON</i>	30	7.19%	11	2.64%	8	1.92%			
<i>Celtic-NF</i>	2	0.48%	1	0.24%	1	0.24%			
<i>Other</i>	0	0	0	0	0	0			
<i>Hybrid total</i>	39	95.1%	83	85.57%	129	48.68%			
<i>Grand total</i>	41	9.83%	97	23.27%	265	63.55%	14	3.36%	

Table 3. The frequencies of Celtic toponyms

Finally, a special group of Celtic place-names have to be mentioned, namely the so-called Hiberno-Norse inversion compounds, which are especially frequent in Cumbria (for an excellent account of such names see Parsons 2011). Their characteristic feature is that they are made up of either two Norse elements or a Celtic and a Norse element, and their order does not follow the Germanic generic + specific configuration, but it is reversed as specific +

generic (see also footnote 4 on page 37). These place-names were likely coined by settlers of Norwegian descent “many of [whom] had come to England from Gaelic speaking areas and following Celtic custom they formed some place-names by reversing the order of the elements” (Cameron 1996: 28).

5.4.1. Celtic generics and specifics

As it has been established, the Brittonic language left a very much unremarkable trace on English, which is also reflected in the toponymic elements that OE borrowed from Celtic. The generics *cumb* ‘valley’ and *torr* ‘craggy hill’ are universally accepted to be loanwords in English from Celtic. This is also evidenced by the fact that both elements occur in monolingual OE formations, such as *Crowcombe* (Somerset) < OE *Crawancumb* ‘valley frequented by crows’ from OE *crāwe* ‘crow’ and *cumb* ‘valley’ or *Worminster* (Somerset) < OE *Wormester*⁴⁶ ‘rocky hill of the snake or dragon’ from OE *wyrm* ‘serpent’ and *torr* ‘rocky hill’. These two generics can co-occur with appellative specifics and personal name specifics too, and are found in highest numbers in the southern counties of England and in the areas near the Welsh border and Cornwall.

Besides *cumb* and *torr*, according to Coates (2007: 181) a small number of Celtic place-name elements also entered the English onomasticon, which means that they were used in toponyms but “rarely attained lexical status”. The possible route for these words into the English language was that OE speakers took Celtic words to be descriptors of topographical features and started using them solely in the function of toponymic elements as none of these items entered the lexicon of the language. The elements in question are the following: **aβon* ‘river’, **cēd* ‘wood’, **cors* ‘bog’, **crüg* ‘barrow’, **eglēs* ‘church’ (from Greek *ecclesia* ‘assembly’), **lux* ‘lake’, **penn* ‘head’, **pull* ‘pool’ and **ros* ‘moor’ (Coates *ibid.*). These words are indicated to be Celtic rather than English in the corpus, and when they occur as part of hybrid forms they are typically coupled with a synonymous explanatory element. They can also surface as simplex names and in a fewer number of instances they are used together with an OE generic.

This configuration means that these words were probably not borrowed into English but are code-switches instead, because of their low frequency and limited context of occurrence. The tautological compounds are in accordance with Coates’ point cited above,

⁴⁶ The ModE form of the name betrays the insertion of an unetymological medial *-n-* probably through folk etymology, under the analogical influence of *minster* as in for instance *Kidderminster* (Worcestershire).

namely that they did not enter the lexicon of the English language, otherwise their meanings would have blocked the addition of synonymous elements. This also shows that the onomastic divorce can enable speakers to create tautological names, which would otherwise go against linguistic and cognitive economy. The fact that the elements were not accepted into the Old English lexicon also hints at the possibility of CS. It should also be kept in mind that place-names can easily be transferred from one speech community to another without speakers necessarily being aware of their meanings, however toponyms are not just mere meaningless labels, but their semantic content is also important as they serve the purpose of identification (cf. the discussion in Section 3.1.).

5.4.2. *Celtic hybrid names*

Apart from the tautological hybrids mentioned in Section 5.4. above, Celtic hybrid toponyms fall mostly into the category of topographical names, followed by names with appellative specifics, and the occasional items with personal name specifics (cf. Table 3 above). In general, Celtic specifics most frequently co-occur with English second elements (196 instances out of 250 hybrid names, which equals 78.4% of all hybrids), followed by Scandinavian generics (49 out of 250, 19.6%), Norman French second elements (4 out of 250, 1.6%) and only one Latin element⁴⁷. With the exception of the category of topographical names, hybrid formations outnumber monolingual ones. It should be reiterated at this point that these figures are only for Celtic place-names found in the territory of England, and thus do not include river-names and names for other geological formations. They do include, however, toponyms created from river-names and from names of other features.

In the category of personal name hybrids Celtic-Scandinavian names stand out most. The corpus indicates the majority of personal names in this section to be of Old Irish origin, which is not surprising given that the largest proportion of these names (14 out of 30, 46.7%) are found in the county of Cumbria in the North West. There are many settlements in this region created or at least influenced by those Scandinavians who came over from Ireland, sometimes also labelled “Hiberno-Norwegians”, and settled in the Northwestern areas of England. The previously discussed inversion compounds also occur most frequently in this region. Cumbria is followed by North Yorkshire (9 out of 30 names, 30%), Lancashire (4

⁴⁷ The sole Latin item, which is probably the personal name *Aust* from *Augustus*, can be found in *Lanercost* (Cumbria) < *Lanrecost* (attested 1169) from Celtic **lannerch* ‘glade, clearing’ and the Latin personal name (Mills 2011: 286).

names) and Lincolnshire (3 names). The other types of Celtic personal name hybrids can be considered to be very rare, even in the smallest category of Celtic toponyms.

In the case of appellative specifics, monolingual formations are much more frequent than in the category of toponyms derived from personal names. There is only one Celtic-Latin hybrid formation in this category, one Celtic-French hybrid⁴⁸, 11 names (representing 13.4% of appellative hybrids) are combined with Old Norse generics and 70 have an English second element (85.36% of appellative hybrids). In this final group, Celtic specifics are coupled with 32 different types of English generics which brings the type-token ratio up to 0.457, meaning that even in such a small sample an extremely wide variety of generics were utilized for coining the names. This entails that OE toponym-formation relied on very precise and specific semantic rules and that even in hybrid names, speakers gave localities names that were not just labels but very nuanced descriptors of the environment regardless of the language of the toponym's first element.

Still in this group, OE *tūn* 'village' is the most frequently encountered generic (16 occurrences, 22.86% of Celtic-OE appellative hybrids), followed by *ceaster* 'Roman camp' (9 instances, 12.86%), with *burh* 'fortification' sometimes in the dative form *byrig* (4 occurrences, 5.71%). These figures are in accordance with the previous discussion of Celtic-English toponyms, namely that they frequently contain *ceaster* and *burh*, which is often accompanied by clipping the original Celtic name. Furthermore, in this class we can find 11 names with explanatory second elements that are the above-mentioned tautological compounds, with generics that include OE *hyll* 'hill', *dūn* 'hill' and *lēah* 'wood, clearing' which were added to Celtic specifics of similar meanings.

Place-names with topographical specifics represent the largest category amongst Celtic names in the corpus, with a combined occurrence of 265 instances (63.7% of all Celtic names), of which there are 120 hybrid ones (28.84%), and 136 (32.7%) monolingual formations. Within the subcategory of topographical hybrids, there are again no Celtic-Latin names, and there is only one Celtic-French item. All eight occurrences of Celtic-Scandinavian place-names with topographical specifics have Celtic river-names as their specifics and in all eight cases the generic is ON *dalr* 'valley'. Amongst the topographical hybrids with English specifics, Celtic river-names show an unassailable dominance with 117 occurrences (97.5%), with the remaining 3 items containing a hill-name as the specific. In comparison, the number

⁴⁸ The Celtic-Latin appellative hybrid is *Lincoln* (Lincolnshire) < *Lindum Colonia* 'Roman colony (for retired legionaries) by the pool' in which the Celtic element is **lindo* 'pool' (Mills 2011: 297). The Celtic-French hybrid is *Cricket* (Somerset) < *Cruchet* (attested in DB) from Celtic *crüg* 'mound' and the French suffix *-et* 'little' (Mills 2011: 138).

of river-name specifics in monolingual topographic names is 67 out of 136, which is still nearly half (49.26%) of the entire class of such names.

The figures outlined above clearly show that linguistically the Celtic population had very little impact on the English toponymic system (and indeed on the English language itself), even though the number of Celtic-related names alone would contradict this. Looking behind the numbers, we can find that personal name hybrids occur in negligible numbers which indicates that on the one hand, in accordance with Celtic place-naming conventions outlined above, personal names do not feature prominently in Brittonic toponyms. On the other hand it also indicates that if landownership changed it changed in a way that an Anglo-Saxon owner took over the estate and if the toponym reflected the original owner's identity it was displaced and substituted with a monolingual English name, and therefore does not occur in the data of the corpus. The Irish-Scandinavian names of Cumbria and other areas of the North represent an entirely different situation and scenario than the place-names carrying Celtic anthroponyms in other parts of England, because the former were created centuries later than the latter which emerged in the aftermath of the Germanic invasion, while the Irish-Norse names emerged during the Viking Age.

The appellative hybrids, that also include tautological names, display the dominance of English generics over Norse and French ones. The low number of names in this group also points in the direction of the negligible effect of Celtic on English place-names. The invading Anglo-Saxons treated the Celtic and Romanized Celtic place-names in different ways depending on the area as well as on the nature and the speed of assimilation (cf. Bölcskei 2012), yet the overarching general trend of limited influence can still be observed. The borrowed elements *cumb* and *torr* discussed above represent the only detectable influence that Celtic had on the system of English generics. In the case of topographical hybrids, the overwhelming majority of instances have a Celtic river-name as the specific. These figures reveal that the other measurable mark the Celtic language left on English toponyms is the names of rivers and other large topographic features that are generally known to show a very high rate of cross-cultural survival (cf. Clark 1992a: 480). In conclusion, all the data presented above indicate the substratal position of Celtic in relation to English.

5.5. The Latin element in English place-names

Despite the linguistic influences that Brittonic and Latin exerted on each other during the Romano-British period, the Romans did not leave a significant mark on the Celtic landscape in terms of toponymy, as only a small number of Roman place-names were coined in Latin and an equally low amount of Celtic forms were modified by Latinization (Rivet 1980: 3) in a way that it was continued to be used in Anglo-Saxon times. This would indicate that a foreign aristocracy ruled over a local population without an accompanying large-scale settlement and peasant migration (like in the case of the Germanic invasion and the Viking conquest), and the proportion of the Latin-speaking population after the Roman conquest was small compared to the number of native Celtic speakers (Rivet & Smith 1982: 11). This state of affairs also entails that the Brittonic speakers maintained their identity under Roman rule and even though in some areas language shift occurred and bilingualism is also likely to have developed, the Celtic population, their language and place-names were not eradicated. As it has already been established, the Romans did not tamper much with Celtic place-names apart from Latinizing their form for documents, and when new toponyms arose, “the name in nearly all cases was Celtic in Latin guise” (Rivet & Smith 1982: 23). The reason behind such a state of affairs could be the difficulty that Celtic resistance posed for the Roman conquest, as mentioned before.

Perhaps the most important historical source of pre-Germanic Latin place-names in Britain is the British section of the *Antonine Itinerary* (Rivet & Jackson 1970, Rivet & Smith 1982: 23-24) known as *Iter Britanniarum*. The Itinerary contains routes along Roman roads in Britannia together with the distances indicated between settlements, the names of which are naturally given in a Latin or Latinized form. This does not mean, however, that these settlements were renamed by the Romans, as explained above, but merely the form of their name was altered to fit into the Latin language Itinerary.

The actual number of genuinely Latin place-names that were continued to be used by speakers of Old English and that survived into Modern English is actually very small, as most Latin toponyms were discontinued and replaced after the Germanic conquest. Many of these names “applied to small places or unimportant features” (Rivet & Smith 1982: 19) which made it unlikely that they would remain in continued use after the end of the Roman period. Only a handful of Latin names survived until the Old English era and until the present-day, which will be discussed below.

Similarly to the number of surviving Latin toponyms, the overall direct influence of Latin on English place-names is in fact barely measurable. In the corpus there are altogether 11 Latin names (0.1% of all the toponyms in the dataset), of which there is only one with a personal name as its specific, and it is also a Latin-OE hybrid⁴⁹. There are two monolingual appellative names⁵⁰, and two appellative hybrids: one Latin-OE (*Templeton* in Devon), and one Latin-ON (*Felixkirk* in N. Yorkshire). Apart from 6 simplex names, there are no other toponyms in the corpus that contain a specific of Latin origin. It should also be noted here that a handful of English-Latin and a couple of French-Latin hybrid names do exist whose second element is Latin, and they will be discussed in their respective subchapters.

Apart from the names just surveyed, an entirely different type Latin toponym elements also occur in the dataset with significantly higher prominence. The overwhelming majority of Latin names in the corpus comes from the Middle Ages and from Medieval Latin (Mills 2011: xv), and function as manorial, distinguishing and commemorative affixes that were added to already existing English, Celtic and Scandinavian toponyms between the 11th and 14th centuries. These additional elements are not counted in the overall number of Latin toponyms because they were not an organic part of the place-names in question but were added decades and often centuries after the names had been created.

The most frequently occurring additional elements are *magna* ‘great’, *parva* ‘little’, *maior* ‘greater’, *regis* ‘of the king’, *episcopus* ‘bishop’, *abbas* ‘abbot’. In all of the cases, the first attestation of the additional elements postdate the completion of the Domesday Book. The practice of adding Latin manorial, distinguishing and commemorative elements was continued even up to recent times, as the case of *Bognor Regis* (West Sussex) illustrates. The original name itself is attested as *Bucganora* c. 975 meaning ‘shore of a woman called *Bucge*’ and *Regis* ‘of the king’ is quite a recent addition from 1929 (Mills 2011: 64).

Altogether there are 110 “affixes” (as they are referred to as by Mills, which will be adopted for the sake of continuity, albeit noting that these are not affixes in the derivational morphological sense) that are added to all types of already existing toponyms. Names affected by these additional elements are only counted once in the corpus, under their respective headings, categorized by the language of the specific. Table 4 below contains the list of additional “affixes” from the corpus along with the frequency of their occurrence.

⁴⁹ *Jacobstowe* (Devon) ‘holy place of St. James (in the Latin form *Jacobus*)’ first attested 1331, from the Latin form of the personal name and OE *stōw* ‘assembly place, holy place’ (Mills 2011: 261).

⁵⁰ *Pontefract* (West Yorkshire) ‘broken bridge’ *Pontefracto* (att. 1090) < Lat. *pons* and *fractus* (Mills 2011: 372). The other name is *Whitchurch* (Shropshire) ‘white church’ that is first attested in 1199 in the Latin form *Album Monasterium* (Mills 2011: 494).

<i>element</i>	<i>meaning</i>	<i>count</i>	<i>percent</i>
<i>magna</i>	great	43	39.09%
<i>parva</i>	little	32	29.09%
<i>regis</i>	of the king	16	14.55%
<i>abbas</i>	abbot	7	6.36%
<i>monachorum</i>	monks	5	4.55%
<i>episcopus</i>	bishop	3	2.73%
<i>abbatissa</i>	abbess	2	1.82%
<i>maior</i>	greater	1	0.91%
<i>castrum</i>	castle	1	0.91%

Table 4. Latin additional elements in the place-name corpus

As it was discussed before, the Latin language did not exert that kind of influence on English that a direct contact situation involving substrates and superstrates would normally exert. Instead, Latin acted as a prestrate to English which means that words were borrowed as a result of the high prestige of Latin especially in the ecclesiastic and scholarly domains. The prestrate nature of Latin is well illustrated by the later additions in Table 4 above, as it can be seen that 5 out of the 9 different types of “affixes” come from the domain of religion and royalty, and the two most frequent ones are equivalents of already existing English words that were probably chosen due to the prestige of the Latin language. While these toponyms on the surface conform to the criterion for hybrid names they cannot be considered to have come about through code-switching.

5.6. The Anglo-Saxon layer

For the sake of full disclosure, and with running the risk of stating the obvious, the starting point of this section should be that by far the largest proportion of names in England is naturally of Anglo-Saxon origin. In the corpus there are altogether 8,117 English toponyms, which make up 78.72% of the entire corpus, in the following break-down. The corpus contains 7,837 (96.55% of all English names) non-hybrid English names, of which 2,384 (29.37%) have a personal name as their specific and 5,276 (65%) have an appellative first element. The number of hybrid names with English specifics and any other generic is 280

(3.45%), and the largest proportion of items in this group is by far taken up by Scandinavian elements of which there are 254 (90.71% of hybrid names). This is not a surprising figure because Old Norse had the most significant impact on English place-names that is on par with its impact on the Old English language in general.

The Scandinavian influence is observable in roughly equal measures in the case of names with English personal name specifics (100 OE-ON hybrids) and with English appellative specifics (152 OE-ON hybrids). These English-Scandinavian hybrid formations will be discussed in greater detail together with the Scandinavian-English hybrids in Section 5.7.2 below. Other hybrid names occur in negligible amounts. There is only one English-Celtic hybrid in the category of personal names and appellatives, there are no English-Latin hybrids in any group, altogether 5 OE-French hybrids, and 19 “other” types which refers to elements coming from languages other than the main four under consideration here (these elements are mostly of a Continental Germanic or of unknown origin). The number of English-Scandinavian hybrids might not seem to be too much compared to English monolingual names, but if they are taken together with Norse-English hybrids (N = 340), yielding altogether 592 Norse-related hybrids, it can be seen that Old Norse was the only language that was able to leave a mark on the English toponymic landscape.

Table 5 below summarizes the frequencies of the various types of English toponyms occurring in the corpus. The line labelled “Scanded” refers to Scandinavianized monolingual English names which were created by sound substitution, and are only counted only once within the group of monolingual names (hence the missing percentages in the Scandinavianized line). These names will also be explored in Section 5.7.2. as part of the analysis of Norse-influenced toponyms.

OE names *N* = 8117 (78.72%)

	Personal name		Appellative		Topographical		Simplex	
	count	percent	count	percent	count	percent	count	percent
<i>Monolingual</i>	2384	29.37%	5276	65%	35	0.43%	142	1.75%
<i>OE-Celtic</i>	1	0.01%	1	0.01%	0	0		
<i>OE-Latin</i>	0	0	0	0	0	0		
<i>OE-ON</i>	100	1.23%	152	1.87%	2	0.02%		
<i>OE-NF</i>	1	0.01%	4	0.05%	0	0		
<i>Other</i>	4	0.05%	15	0.18%	0	0		
<i>Scanded</i>	24		100					
<i>Hybrid total</i>	106	4.26%	172	3.16%	2	5.71%		
<i>Grand total</i>	2490	30.68%	5448	67.12%	35	0.43%	142	1.75%

Table 5. The frequencies of Old English toponyms

Concerning monolingual formations, the number of place-names containing appellational specifics is more than the double (2.2 times greater) of the number of toponyms with personal name specifics. This indicates that Anglo-Saxon place-names tend to encapsulate more the environmental features and other attributes of settlements than the names of owners and overlords, meaning that when coining the names, OE name-givers relied more heavily on the lexicon of the language for specifics than on personal names. This in turn reveals that toponyms are in fact meaningful elements of language because of the minute differences of meanings they cover and reflect, and are not just simple labels. It can be argued that generics represent “cognates” of their free lexical equivalents (cf. Colman 1992, Hough 2012a) because some of them never occur outside place-names, therefore they are different from free equivalents even though they carry the same meaning. I do not necessarily agree with this terminology, because *cognate* would imply that the two words share the same etymon but come from different languages, therefore I would prefer the term “positional variant” (as in allophones and allomorphs) or on this analogy “onomastic variant” instead. Specifics on the other hand are always drawn from the pool of free words available in the lexicon of the language.

5.6.1. Old English generics and specifics

In the corpus of settlement names, there are altogether 404 different types of English generics, the majority of which are hapax legomena, i.e. they occur only once in the entire corpus. The most frequent element is *ūn* ‘enclosure, village’ with 1445 occurrences, followed by *lēah* ‘clearing’ (614 instances) and *hām* ‘homestead’ with 445 occurrences⁵¹. As it can be seen even from the first three items, the generics show a very uneven distribution, with the outstanding frequency of the first element and the sudden decrease in frequency of all following items. Table 6 below presents the fifteen most frequent English generics in the corpus. It should be kept in mind that these counts are to be taken for the entirety of the corpus, including generics found in Norse-English hybrid names, which brings the total count of names with English generics up to 8653. The percentages given in the table are calculated from this higher number. The meanings are given in accordance with Mills (2011).

⁵¹ Out of these 445 occurrences 155 show uncertainty and hesitation between *hām* ‘homestead’ and *hamm* ‘enclosure’ but are counted within the total tally for *hām*.

<i>generic</i>	<i>meaning</i>	<i>count</i>	<i>percent</i>
<i>tūn</i>	enclosure, farmstead, village	1445	16.7%
<i>lēah</i>	wood, clearing	614	7.09%
<i>hām</i>	homestead	445	5.14%
<i>ford</i>	ford	328	3.8%
<i>dūn</i>	(low) hill	255	2.95%
<i>feld</i>	field	184	2.13%
<i>worth</i>	enclosure	182	2.1%
<i>wella</i>	spring, stream	155	1.79%
<i>halh</i>	nook, corner	141	1.63%
<i>burna</i>	stream	127	1.3%
<i>denu</i>	valley	122	1.47%
<i>burh (byrig)</i>	stronghold	113	1.31%
<i>ēg</i>	island	110	1.27%
<i>hyll</i>	hill	101	1.17%
<i>cumb</i>	valley	77	0.89%

Table 6. The distribution of the most frequent English generics

As it can be seen from the table, the generics show a very sharp decline in the first four instances after which the decline in numbers shows an ever gradual tendency of slowing down, leading towards the appearance of hapax legomena. This situation is characteristic of long-tailed distributions whereby the first few items show outstanding frequencies that gradually decline. This is evident from the dataset above in which the first four generics show very high frequencies and account for nearly one third of all the English generics in the corpus (2759 occurrences, 31.9% of OE generics). This state of affairs means that apart from the most frequently occurring and most general and widely applicable generics, in OE there was a huge number of generics with very narrow and specialized meanings that could be used on a one-off basis for the description of settlements.

Cameron (1996: 66-72) considers place-names in *-ingas* and *-ingahām* to be the oldest Anglo-Saxon place-name formations. This generic goes back to the plural form of *-ing* which was added to a personal name to form a patronymic name, and in the plural it denoted dynasty, and later the dependents or people of a person. The names, according to Cameron (ibid.) represent “a social organization which must have preceded the establishment of

kingdoms” and such names are originally not place-names but group names (cf. also Fellows-Jensen 1978b: 26). Jones (1990: 53) also concludes that these names must be amongst the earliest ones, but also notes that they could well have been formed during a phase of internal colonization when the Anglo-Saxons settled on virgin land or took over already existing estates. Jones’ (ibid.) other candidate for earliest place-name elements is OE *wīc*⁵² which comes from Latin *vicus* ‘vill’ which would indicate a continuity between the Romano-British and the Anglo-Saxon populations. It has also been proposed that names in *-hām*, *-inga*, and *-ingas* indicate early Anglo-Saxon settlement and point to a continuity with Romano-British settlements and their Germanic appropriation, but the archaeological evidence cannot prove the latter claim (Fellows-Jensen 1978b: 24-25). The *-ing-* element occurs almost exclusively with *tūn* ‘village’, while *-inga-* occurs mostly with *hām* ‘homestead’ but there are also a number of generics that it can co-occur with. The collocational patterns and frequencies of *-ing-* and *-inga-* are shown in Table 7 below for the set of toponyms with personal name specifics (N = 728). On the other hand, names in *-ing* and *-inga(s)* are coupled much less frequently with appellative specifics (N = 75), which is the expected arrangement given that its original function was to indicate the dependencies of a certain person.

<i>first pt</i>	<i>second pt</i>	<i>count</i>	<i>percent</i>	<i>meaning</i>
<i>ing</i>	<i>tūn</i>	334	45.88%	enclosure, village
<i>ingas</i>		101	13.87%	
<i>inga</i>	<i>hām</i>	90	12.36%	homestead
<i>inga</i>	<i>lēah</i>	12	1.65%	wood, clearing
<i>ing</i>		8	2.13%	
<i>ing</i>	<i>cot</i>	8	2.13%	cottage
<i>inga</i>	<i>halh</i>	5	0.68%	nook, corner
<i>ing</i>	<i>dūn</i>	5	0.68%	(low) hill
<i>ing</i>	<i>denn</i>	3	0.41%	
<i>ing</i>	<i>hām</i>	3	0.41%	homestead

Table 7. The frequencies of anthroponymic names in *-ing* and *-inga(s)*

⁵² Mills (2011: 525) gives the generic’s meaning as follows: “earlier Romano-British settlement; dwelling, specialized farm or building, dairy farm; trading or industrial settlement, harbor.”

As it can be seen from the table, the familiar long-tailed distributional pattern emerges again, but this time the drop is very steep between the first three elements, meaning that the generics *ingtūn* and *ingahām* could even be considered to be set phrases or deeply entrenched compounds with a very high collocational force between their constituents. There are 9 other types of second elements that occur after *-ing* with frequencies ranging between 1 and 3, so altogether there are 17 different types of generics that follow *-ing* in the corpus giving a very low type-token ratio of 0.03 which indicates a lack of variation and diversity in these generics. On the other hand, second elements following *-inga(s)* show much more variation, as apart from the generics given in the table above, *-inga(s)* co-occurs with 97 second elements, all of which are hapax legomena.

Specifics on the other hand show a more even distribution by which I mean that most specifics occur only once in the corpus and the most frequent ones are not represented in more than a handful of instances. Coupled with the fact that generics show a somewhat equal diversity albeit with more high-frequency items, this indicates that English place-names (and by extrapolation probably place-names in general) are semantically highly loaded and capable of expressing minute distinctions in meaning.

5.7. The Norse layer

The two major types of toponyms of Scandinavian origin which occur in the Danelaw area and the North West are purely Scandinavian names, containing only Old Norse elements (either from the Eastern or Western variety) and hybrid names, which are created by cognate substitution, Scandinavianized through phoneme substitution or created as an originally hybrid name, however this last option can be subject to debate. This section surveys the possible origins of the English-Scandinavian hybrid toponyms as well as the purely Scandinavian settlement names.

After the Treaty of Wedmore between Alfred and Guthrum, the Danes settled in the Danelaw area and then gradually assimilated into the English society and shifted to the use of the English language. The exact nature of the settlement can be subject to debate, as to whether a massive, coordinated migration took place with a wave of peasant migration following the settlement of the warriors of the great army (Fellows-Jensen 1968: xxii-xxiii, Cameron 1996: 75-76, Hadley 2000: 19-20), or whether only a relatively small group of high-ranking Viking soldiers settled, who took control of the north-eastern part of the land and then

subsequently assimilated into the native population (Sawyer 1962)⁵³. The place-name material seems to support the former setup, and the idea of a significant Norse presence in and Norse migration to England (cf. also Townend 2014: 95, 98), especially the fact that the most frequently occurring type of Scandinavian place-name is a habitative name⁵⁴, which combines a Norse personal name and a Norse generic (547 instances, or 51.8% out of 1056 place-names with personal name specifics), which would lead one to believe that a rather large number of settlers migrated to this area and they established their settlements on hitherto uninhabited pieces of land. Whereas those hybrid place-names (the so-called Grimston-hybrids), which contain a Scandinavian personal name and an English generic would suggest the presence of a Scandinavian overlord or owner of an otherwise English settlement. It seems likely that Grimston-hybrids came about when Scandinavian invaders took over English settlements and the English generic was kept, but the specific was substituted with the name of the new owner (Fellows-Jensen 1985: 180, Reaney 1987: 170-171, Cameron 1996: 75). Those place-names, however, which are made up of an English personal name and a Scandinavian generic could reflect the exact opposite of this configuration, namely English dominance over a Scandinavian settlement, or, as will be shown later, that English speakers coined these names using a Scandinavian generic.

Furthermore, appellative hybrids (or Carlton-hybrids), which are made up of common nouns are also quite frequent (265, or 30.85% out of 859 appellative place-names), but not nearly as much as pure Scandinavian formations (419, or 48.78% out of 859⁵⁵). If the pieces of evidence which place-names can provide are taken together, then it may be assumed that

- (i) it is unlikely that only soldiers of the *micel here* settled in the area of the Danelaw,
- (ii) the amount and types of Norse place-names points to extensive and significant settlement by the Scandinavians, which is reflected by the dominance of habitative names, especially those in *by*⁵⁶ ‘village’ and *þorp*⁵⁷ ‘secondary settlement’, ‘outlying farmstead’, with purely Scandinavian names frequently appearing on agriculturally less attractive sites (cf. Clark 1992a: 485), meaning that the new settlers established their settlements on empty pieces of land,

⁵³ For recent synthesizing discussions of the issue see Moskowich 2012: 32-36, and Townend 2014: 95-112.

⁵⁴ Habitative names are defined here as names “which originally denoted some structure or structures used for habitation, shelter or other purposes by man or animal” (Fellows-Jensen 1978: 136).

⁵⁵ See *Table 1* below for further figures.

⁵⁶ 585 (48.79%) occurrences out of 1199 tokens of Scandinavian generics in SSNY, SSNEM, SSNNW and Mills (1998), see *Table 2* below for further figures.

⁵⁷ 273 (22.77%) occurrences out of 1199 tokens *ibid*, see *Table 2* below for further figures

- (iii) Grimston-hybrids (or personal name hybrids) indicate settlements with mixed population, or the very least a village that has quite probably undergone a change of ownership, and
- (iv) hybrid appellative place-name formations and Scandinavianized names created via the substitution of cognate words reveal a degree of mutual intelligibility and mixing between the two languages, especially in those cases where the various attested forms show hesitation and the interchangeable use of English and Scandinavian elements (cf. Fellows-Jensen 1972: 112, 137, Townend 2002: 60⁵⁸).

These place-names, however, shall be handled with care, as Hadley (2000: 332-333) also calls attention to the fact that “it is not appropriate to make a simple connection between Scandinavian place-names and places of Scandinavian settlement”. Indeed, due to the prevalence and outstanding frequency of place-names in *bý* and *þorp*, these generics could have easily found their way into English (cf. Janzén 1972: 8) and have become productive in the language, therefore the role of analogy in place-name formation shall not be underestimated either. It is very likely that English speakers gave to their outlying settlements names in *þorp* (Cameron 1996: 80), and the place-name evidence also seems to support this assumption. There are altogether 198 instances of place-names in *þorp* which have a personal name specific, 51 of these (25.75%) have an English personal name and 130 (65.65%) have a Scandinavian personal name as their first element. In contrast, out of 362 settlement names in *bý*, 291 (80.38%) have Scandinavian personal name specifics and only 34 (9.39%) have English personal name specifics, and there are only 37 (10.23%) appellative place-names in this group. There is also a great number of names which are new formations and are first attested after the conclusion of the Domesday survey, sometimes as late as the 13th century, by which time direct contact with the Scandinavians had long ceased to exist. Furthermore, Scandinavian personal names were not exclusively borne by people of Scandinavian descent, but they were on many occasions adopted by the English, much the same way as Norman personal names were adopted after the Norman Conquest. Therefore, either group of speakers could have engaged in the creation of those settlement names which contain personal names of Scandinavian origin.

In summary the following four major types of Norse-related place-name formation can be distinguished:

⁵⁸ Townend (2002: 57-59) describes 192 instances of clear cognate substitution in Scandinavianized appellative place-names, which is definitely not a negligible figure, and it goes to show that this was not a marginal phenomenon and that the Vikings were able to understand the place-names they encountered.

- (i) pure Scandinavian formations
- (ii) personal name hybrids (Grimston-hybrids)
- (iii) appellative hybrids (Carlton-hybrids)
- (iv) Scandinavianized place-names

From the survey and analysis of the collected place-name data it has emerged that Scandinavian specifics combined with Scandinavian generics is the most frequently occurring type both in the category of settlement names with personal name specifics (51%), and non-personal name (or appellative) specifics (49%). Table 1 summarizes the distribution of place-names with personal name specifics (N = 1056), and the distribution of place-names with non-personal name specifics (N = 859).

Some commentary concerning specific cases and lines of Table 1 below is due here. Owing to the fact that category names such as *Grimston-hybrid*, ‘Grímr’s village’ (SSNY 128, Ekwall 1980: 206, Mills 1998: 157) and *Carlton-hybrid*, ‘village of the freemen or peasants’ (SSNEM 183, Ekwall 1980: 87-88, Mills 1998: 71) entail a sense of exclusivity in favor of a Scandinavian + English order and origin of elements, I propose the use of a more neutral and broader *personal name hybrids* and *appellative hybrids*, respectively, which can also include an English + Scandinavian order and origin. Cameron (1996: 74-75) calls attention to the fact that not all places with the name *Grimston* follow the same pattern, so he proposes that hybrids specifically made up of an ON personal and the OE generic *tūn* be collectively called *Toton*-hybrids (‘Tovi’s village, estate’).

The E/N label in Table 1 below stands for elements which could be either Old English or Old Norse: dubious cases in which the origin of the first element cannot be determined with certainty on the basis of the attested forms. For instance *Autby* (attested forms in the Domesday Book include *Aluuoldebi*, *Alwoldebi* and *Alduluebi*), a depopulated settlement which contains either the OE personal name *Ælfweald* or the (rare) ON personal name *Alfvaldr* (Fellows-Jensen 1978a: 34) or the appellative *Bowthorpe* (attested Domesday Book forms include *Bergestorp* and *Buretorp*), a lost settlement which contains either OE *beorg* or ON *berg* ‘hill’ (Fellows-Jensen 1978a: 104).

Scandinavianized names are defined as toponyms containing an English specific and an English generic the first element of which has undergone phoneme substitution in order to accommodate the English name to the Scandinavian phonological system, for instance the previously mentioned *Fiskerton* and *Digby*. In the case of Table 8, the ‘Other’ label refers mostly to continental Germanic names.

<i>ON names</i> <i>N = 1382 (13.4%)</i>		Personal name		Appellative		Other		Simplex	
	count	percent	count	percent	count	percent	count	percent	
<i>Monolingual</i>	547	39.58%	446	32.27%	5	0.36%	18	1.3%	
<i>ON-Celtic</i>	0	0	0	0	0	0			
<i>ON-Latin</i>	0	0	0	0	0	0			
<i>ON-OE</i>	214	15.48%	126	9.12%	2	0.14%			
<i>ON-NF</i>	0	0	0	0	0	0			
<i>Other</i>	19	1.37%	5	0.36%	0	0			
<i>Hybrid total</i>	233	29.87%	172	29.81%	2	28.57%			
<i>Grand total</i>	780	56.44%	577	41.75%	7	0.51%	18	1.3%	

Table 8. Distribution of Old Norse names in the corpus

The data in Table 8 show the unassailable dominance of purely Scandinavian formations, which, as discussed in the previous sections, hints at rather extensive settlement. It can also be seen from Table 8 on OE names that appellative names contain a significantly higher proportion of Scandinavianized items, which seems to be in accordance with the presupposition that personal names are less likely to undergo Scandinavianization or Anglicization.

5.7.1. Scandinavian hybrid names

In the cases of Anglo-Scandinavian hybrid place-names it cannot be established with certainty which language functions as the receiving language, i.e. the one which broadly speaking supplies the overall structure, and which functions as the language whose forms are inserted into the structure. It is similarly difficult, or even impossible, to declare in certain hybrid names whether they are the result of cognate substitution or they were formed by an “Anglo-Scandinavian” population (Fellows-Jensen 1978a: 205). This is in accordance with the characteristics of congruent lexicalization, namely that it “often involves **bidirectional code-mixing** [emphasis original] since there is no matrix language” (Muysken 2000: 132). It can be seen from the data that there are 126 appellative hybrids (14.67% of all such formations) which have an ON + OE order and origin of elements, while 152 (17.69%) appellative hybrids follow the opposite order. This means that the number of ON + OE and OE + ON hybrids are roughly the same, so there is indeed no dominant matrix language. The shared linguistic structure and the other similarities existing between the languages facilitate congruent lexicalization and bring about bidirectional mixing.

As it was described in Section 2.2.3 it is possible to analyze hybrid compounds within the framework of congruent lexicalization. Muysken (2000: 151) cites the bidirectionality of coining hybrid compounds as evidence for congruent lexicalization: in his sample of English-German hybrids both English-German and German-English compounds occur. Muysken (ibid.) also notes that these hybrid compounds are “one step away from loan translations” whereby both elements would be replaced by words from the other language. The fact that English-Scandinavian and Scandinavian-English hybrid toponyms occur in almost equal numbers in my corpus and the high incidence of uncertain elements (107 for personal name specifics and 62 for appellative specifics) also points in the direction that they arose through code-mixing of the congruent lexicalization type.

Many of the Scandinavian and Scandinavianized place-names are first attested only from 1086, in the records of the *Domesday Book*, therefore there is no reliable data available about the date when these place-names could have been first used, and there is no evidence of whether they were transformed from English names or were new coinages. It is very likely, however, that hybrid names are later formations than purely Scandinavian ones, which were created after the initial phase of settlement, whereas hybrid ones came about in a period when population mixing, sociopolitical adjustment and assimilation was taking place (cf. Fellows-Jensen 1995a: 58).

It can also be argued that the creation of hybrid names and names containing cognate substitution both resulted in the production of mixed names, which are made up of elements from two different languages, therefore constitute instances of code-mixing. This situation, however, cannot be classified as borrowing with certainty. Borrowed elements can either fill lexical or semantic gaps, oust native terms or, in some cases, lead to the continued use of native items but with a restricted or modified meaning or as part of a different stylistic register, rather than be used in such a fashion that allows their coexistence alongside native terms without semantic change or stylistic modification occurring in either of the affected words. In our case, quite a vast array of different appellative specifics are used, most of them rather infrequently, and – due to the fact that none of these elements displaced any of the English ones – the situation at hand is more likely to be classified as code-mixing than borrowing.

Gammeltoft (2007: 481) argues that

“[w]hen a place-name or a place-name element has been borrowed from one language to another, it becomes part of the borrowing language’s onomasticon. From then on it may be used to coin new names in that

language. Thus, when a place-name appears to be a hybrid of elements from two languages, it is in reality not. It is a monolingual coinage utilising a borrowed place-name or a borrowed place-name element.”

This explanation of ‘pseudo-hybrid’ formation through borrowing can definitely be true for hybrid names in *bý*, *þorp* and *tūn* all of which have a very high frequency of occurrence, making borrowing a likely explanation. However, it seems rather unfeasible that such a huge number of diverse elements, both specifics and generics, all capable of describing minute distinctions in meaning were borrowed into either Old English or Old Norse to create monolingual hybrids, while maintaining their subtle distinctions of meaning. It seems much more likely that these formations are indeed true hybrid ones.

Table 9 below shows the distribution of the first fifteen most frequently occurring tokens of Old Norse (N = 1199) generics. The meanings of Old Norse and Old English generics given here are based on Fellows-Jensen (1972, 1978a, 1985), Ekwall (1980) and Mills (1998, 2011). The frequency counts for the generics presented here are the combined numbers of generics with personal names and appellatives, as well as monolingual and hybrid formations.

<i>generic</i>	<i>meaning</i>	<i>count</i>	<i>percent</i>
<i>bý</i>	village	585	48.79%
<i>þorp</i>	secondary settlement	273	22.77%
<i>holmr</i>	water-meadow	43	3.59%
<i>dalr</i>	valley	42	3.50%
<i>þveit</i>	clearing	38	3.17%
<i>toft</i>	curtilage	30	2.50%
<i>lundr</i>	grove	26	2.17%
<i>haugr</i>	hill	23	1.92%
<i>bekkr</i>	beck	20	1.67%
<i>skáli</i>	hut	17	1.42%
<i>vað</i>	ford	16	1.33%
<i>viðr</i>	wood	14	1.17%
<i>berg</i>	hill	12	1.00%
<i>kjarr</i>	marsh	10	0.83%
<i>skógr</i>	wood	10	0.83%

Table 9. Distribution of Old Norse generics

Similar long-tailed patterns can be observed in the case of Scandinavian generics as could have been observed within the OE dataset above. This again reveals that Scandinavian settlers did not rely only on a couple of very frequent generics, but actually utilized a very wide range of available generics. This situation is in contrast with the Celtic generics in Celtic hybrid names and the Latin additional elements discussed before, as these latter two types of items only occur in low frequencies and with low token counts.

Old English and Old Norse generics, on the other hand, show less overlap than the specifics, which could have led to the borrowing of the frequently occurring elements for which no native English term was available, e.g. *toft* ‘homestead, curtilage’, *holmr* ‘water meadow’, *þveit* ‘clearing’. It should also be emphasized that borrowing is not something that happens abruptly and instantaneously, but it is the terminal point of a gradual process, which can, and quite probably does, include code-switching and/or code-mixing, and which is rather closely related to frequency of use, because if a foreign element is used infrequently in a given language then it is much more likely to be an isolated case of code-switching.

It can be seen from the data above that habitative generics, and thus habitative names, seem to be dominant both in the case of Norse and English. The Norse material shows similar, long-tailed distribution of frequencies as does the group of English generics, and, more importantly, it shows a great variety of different elements. This means that settlement names created using these items cannot reflect the use of a few, fossilized generics, but rather reflect that they were created utilizing a wide repertoire of productive and actively used elements capable of capturing minute differences in meaning.

Concerning the data, it is revealed by the analysis that 20 different generics occur in the group of English + Scandinavian (OE + ON) appellative hybrids and 37 different generics are to be found in the Scandinavian + English (ON + OE) group of settlement names. The type-token ratio of generics in the case of English-Scandinavian hybrids is 0.131 with 20 generic types and in the group of Scandinavian-English hybrids the ratio is 0.293 with 37 generic types. Due to the higher degree of heterogeneity in S+E toponyms it can be assumed that those names were created by a Scandinavian population either through cognate substitution or as new formations through using English generics. In the first group *bý* (55 occurrences out of 152, 36.18%) and *þorp* (28 instances, or 18.42%) are by far the most frequently occurring elements, followed by *bekkr* (8 instances, or 5.26%), while in the other group *tūn* has the highest frequency (40 out of 126, 31.74%), followed by *lēah* (5 instances, or 3.96%) and *ford* (5 occurrences, or 3.96%). There is an observable difference between these two types, which is most clearly reflected by the fact that the top two generics in the E+S

group account for more than half (54.6%) of all the token occurrences, while in the S+E group there is only one element, which has an outstanding frequency, and the other ones show more sporadic distribution. This means that it is unlikely that such a great array of different English generics were borrowed by the Scandinavians, and vice versa. It is possible that *bý* and *þorp* were accepted into English where they were used productively, as discussed before, and that OE *tūn* made its way into Scandinavian, where it became productive. It is also likely that English appellative names in *bý* and *þorp* are young formations, as well as Scandinavian appellative names in *tūn*, given that *tūn* was a very productive English generic (Ekwall 1980: xiv-xv), and *bý* and *þorp* were highly productive Scandinavian ones.

It should also be mentioned that while ON *þorp* had its cognate in Old English as *þrop*, its incidence is far lower than its Scandinavian counterpart, and occurs very infrequently in purely English, non-hybrid settlement names outside the Danelaw area, e.g. the simplex name *Thrupp* < OE *þrop* ‘outlying farmstead or hamlet’ in Gloucestershire, so much so that Mills (2011) cites only a handful of examples for its occurrence, most of which come from regions that were subject to Scandinavian influence, such as

- (i) *Throphill* (Northumbria, first attested in 1166 as *Trophill*) ‘hamlet hill’, from OE *þrop* ‘outlying farmstead, hamlet’ and OE *hyll* ‘hill’, in which case the element is used in the function of the specific (Mills 2011: 428) and
- (ii) *Abthorpe* (Northamptonshire, first attested in 1190 as *Abetrop*), ‘outlying farmstead or hamlet of a man called *Abba*’, from an OE personal name and either the OE generic *þrop* or the ON generic *þorp* (Mills 2011: 3).

Furthermore, OE *tūn* also had its cognate in Old Norse as *tún*, but as Fellows-Jensen (1978a: 175-176) has demonstrated, its occurrence in Denmark is very infrequent and the element was quite probably unproductive at the time of the Viking conquest of England, therefore it is unlikely that this element was either transplanted by the Viking settlers to England or that it played a significant role, or any role at all, in the facilitation of the emergence of Scandinavian-English hybrid place-names in *-tūn* in England. However, as Fellows-Jensen (*ibid.*) also notes, this generic occurred rather frequently in Norway, Iceland and central Sweden but with a different meaning and borne by different types of settlements than in England or Denmark. Since the vast majority of Scandinavian settlers in England, especially in Yorkshire and the East Midlands, came from Denmark, the strongest piece of evidence against the conflation of OE *tūn* and ON *tún* seems to be the fact that the latter had already obsolesced by the time the Danes came into any sort of contact with the English.

The English-Scandinavian hybrid names treated in this dissertation could be analogical formations based on existing ON + ON and OE + OE names, respectively or it could also be the case that they were created through the active and productive use of the generics. In both groups a rather heterogeneous selection of specifics can be found, with a predominance of adjectives, and nouns denoting humans. The code-mixing found in those hybrid settlement names which contain a Scandinavian specific and an English generic likely arose in the process of the Norse population's assimilation and shift to English, and was probably necessitated by the situation itself. The emergence of most of the mixed place-names discussed above cannot be satisfactorily explained as borrowing, because we have two Norse generics and one English generic with outstandingly high frequency of occurrence and 18 and 36 more types of generics from Old Norse and Old English, respectively, with significantly lower frequencies. Nor can they be explained in a satisfactory way as being young mixed formations, due to the gaping holes in textual evidence regarding settlement names antedating both the Viking and the Norman French invasion, therefore it is proposed that the hybrid names can be labeled instances of intralexemic (or word-internal) cognate substitution ("non-constituent mixing", Muysken 2000: 129, 137), which might form a broader subtype of historical code-mixing, more specifically of congruent lexicalization. In this case, there is a mixed compound word, which is a complex lexeme, the overall structure of which is shared by both languages, because Old English and Old Norse shared their place-name formational practices, which meant a specific + generic order of elements and the use of the dative plural and genitive cases. For word-internal cognate substitution, a preexisting form is required, which in our case is supplied by English, into which forms from a closely related cognate language can be inserted thereby creating cognate substitution and a mixed-language lexical item.

5.8. The Norman French layer

As it has already been discussed in Section 3.3. the Norman French conquerors (along with their various allies) displaced much of the English upper classes, and were considerably fewer in number than the English-speaking population. Due to these two factors the occurrence of French-derived, French related or French influenced place-names in England is rather scarce (Cameron 1996: 20). However, owing to the wholesale replacement of English nobility, the Norman French language was clearly in a superstratal position. The situation between Old English and Norman French (at least in the case of toponyms) has been likened to

the one that existed between Celtic and Latin in Romano-British times, based on the similarities of the presence of a colonial elite and their negligible influence on toponymy (cf. Rivet 1980, see also *ibid.* for a critique of drawing such parallels).

One of the differences though between the OE-NF and the Celtic-Latin situation is that Old English and Old Norse personal names were practically eradicated from the language after the Norman Conquest Hough (2000: 4) and replaced by Biblical names and continental Germanic names through Norman influence. Also, many French loanwords were introduced but a considerable proportion of the OE and ON vocabulary also survived. The relationship between English and Norman French is a typical example of superstratal influence with massive lexical borrowing and very limited effect on toponyms, as only a few dozen French names and French hybrids emerged and no French place-name elements were borrowed into English.

Altogether there are 57 French-related toponyms in the corpus (0.55% of all the names) which makes it the second smallest category after Latin names, but similarly to Latin elements, a number of English place-names contains an additional manorial “affix” of French origin which is always a personal name or a family name in reference to the owner of the given estate. There are 18 place-names with personal-name specifics (31.58%), all of which are hybrids: 8 French-English ones and 10 French-Old Norse ones. In the category of appellative specifics, the corpus contains 35 names (61.4% of French names) the majority of which (N = 29) are monolingual formations and the remaining 6 toponyms are French-English hybrids. Finally, the corpus contains 4 simplex French names.

As it can be seen from the numbers above and from the socio-historical background discussed in the previous chapter, Norman French influence on English place-names is very much negligible. With the exception of 7 names, all instances of French-related toponyms are first attested after the completion of the Domesday Book, mostly from the period between the 12th and 14th centuries with a pronounced imbalance in favor of names coined (or at least attested) in the 12th century.

VI. Conclusions

6.1. Overview of results and concluding remarks

In this dissertation I analyzed place-name formations in England, with special attention to hybrid ones, which emerged as a result of the language contact situations that English was a participant of during its development since the mid-5th century. I wished to argue that the English-Scandinavian hybrid place-names can be classified as instances of word internal (or intralexemic) code-mixing of the congruent lexicalization type, and in the cases of Scandinavianized names can be seen as instances of intralexemic cognate substitution. Furthermore, my aim was also to be able to ascertain whether or not the stratal relationship of languages play any role in the emergence of hybrid toponyms and whether or not these names can in fact be classified as instances of CM. This dissertation did not intend to go anywhere beyond the realm of language, therefore it did not aim to make any claims or draw any conclusions regarding the distribution of Scandinavian and other layers of settlement based on the place-names alone, which, as the author is convinced, cannot be viewed as the sole indicator of such distribution.

From the corpus analysis presented in the previous chapter, it emerges that apart from Old Norse, no other language exerted significant influence on the toponymic system of (Old) English. Pre-Celtic names are restricted to river-names, and to names denoting large topographical features, such as hills, fields and woods. Such elements are mostly conserved in the toponyms of later layers of settlers and languages. Celtic names represent an important layer in the domain of English place-names. The Celtic influence is observable mostly in toponyms derived from river-names and hill-names and also from larger topographical features (much like in the case of the pre-Celtic layer). The Latin elements in the corpus are restricted to a handful of actual Latin-derived toponyms and to a larger group of manorial, distinguishing and commemorative “affixes” that were superadded to already existing Celtic, English and Norse place-names long after the conclusion of the DB survey. The Norman French language also had very limited influence on English toponymy.

Naturally, the Anglo-Saxon elements constitute the largest group of specifics and generics within the corpus of English settlement names, which are followed by names of Old Norse origin. In general it can be said that a wide variety of different generics were used both in the case of OE and ON names, which means that a substantial Scandinavian population inhabited areas of England, mostly in the north east.

Concerning the genetic and stratal relationship of languages (as per the fourth research question⁵⁹ from Chapter 1) it can be stated that these factors significantly influence the outcome of linguistic interference in the domain of toponymy. Substratal languages (such as pre-Celtic and Celtic in our case) typically do not exert any significant influence on toponyms and only form hybrid names out of necessity and through the unavoidable use of their hydronyms and toponyms by newcomer settlers. The overwhelming majority of Celtic-English hybrid place-names contains a Celtic river-name as their specific combined with an English generic. The superstratal Norman French language was also unable to leave a measurable mark on place-names in the corpus, which resulted from the absence of large-scale settlement by speakers of French. Only the adstratal Old Norse was able to affect the OE toponyms to a considerable degree, with frequently occurring ON monolingual names, followed by hybrid place-names and Scandinavianized monolingual OE names. All the languages that (Old) English got into contact with during its history and are represented in the corpus are genetic relatives of English. However, only Old Norse comes from the same Germanic branch of IE as English (albeit OE is a West Germanic language, while ON is a North Germanic one), which facilitated hybridization, along with the wave of Scandinavian peasant migration and settlement in England.

To answer the research question cited above, it can be said that only an adstratal relationship between two languages, preferably of close genetic relationship and accompanied by settlement, will yield toponymic influence. Substratal languages only provide toponyms that new settlers need to utilize when accommodating to their new environment. Hydronyms are the best candidates for substratal names that survive language shift and even language death. Superstratal influences on local toponyms are usually negligible if the introduction of the superstratal elite is not accompanied by massive settlement and migration of the speakers of that language. Finally, prestrates only influence toponyms by virtue of their prestige and by providing additional elements to be used for distinguishing from each other those settlements that bear the same name and for dedicating settlements to certain people. Lexical elements from the prestrate is typically found in certain domains (ones in which the prestrate culture exhibits some form of cultural dominance), and will therefore mostly provide lexical items for additional elements coming from those domains.

⁵⁹ Research questions 1-3 have already been answered by the chapters on the theoretical contextualization of the main topics of the dissertation. For the sake of brevity, they will not be recapitulated here, but I will kindly refer the reader to Chapters II and III.

The remainder of the names that are not discussed in the previous chapter fall under the category of “uncertain names”, which mostly includes toponyms of Anglo-Scandinavian origin. In these cases it cannot be determined with certainty which language supplies the given form found in the place-name. This can be taken to be indicative of the presence of homophonous diamorphs which results from the close genetic relatedness of the languages and which facilitates the creation of hybrid names via congruent lexicalization.

Concerning the fifth research question, historical code-mixing and borrowing is represented in the various layers of toponyms in different ways. The wide variety of ON generics hints at code-mixing especially in the case of infrequent generics and hapaxes. These types of elements can also be found in much smaller quantities in other layers of toponyms in the corpus, which can also indicate that they are the results of CM. Based on the frequency counts of generics in the corpus, high-frequency generics are very likely to be borrowings, and low-frequency ones are likely candidates for CM. Celtic, Latin and French elements can become morphologically integrated into English, and they are also affected by all the regular changes and onomastic sound changes (resulting mostly from the onomastic divorce) that affect all the other toponyms as well.

Apart from a few Scandinavian hybrid names containing Continental Germanic elements (only specifics occur from CGmc, no generics, 24 in total out of 1382 ON-related names), English is the only language that supplied generics for ON hybrid names, and conversely, ON is the only language that participated in the creation of English hybrid toponyms to any considerable degree. English-Scandinavian and Scandinavian-English hybrid names show the signs of bidirectionality, morphological integration, word-internal mixing and the presence of homophonous diamorphs, all of which are key features of congruent lexicalization. It is also highly possible that code-mixing in ON-OE and OE-ON hybrid place-names constituted an unmarked choice hence their outstanding frequencies and continued survival. On the other hand, Celtic and French hybrid names exhibit unidirectionality, a key feature of insertional CM. With the exception of post-DB additional distinguishing elements, Latin names occur in such low numbers that the results are practically inconclusive, and reveal that Latin did not play a significant role in the formation of hybrid toponyms with the English language.

Concerning the sixth, and final, research question, the types of code-mixing described in Chapter II are indeed observable in hybrid place-names in England, as discussed in the previous two paragraphs. The hybridization processes described in Chapter III are also manifested in the toponymic data of the corpus. In the case of OE-ON hybrids, the affected

names were created as descriptive lexical items with relevant semantic content that both parties were able to decode and served functional purposes (cf. Hough 2012a). The purpose, the quality and the quantity of English-Scandinavian hybrid names is significantly different from all other types of hybrid names in England. The Celtic hybrids (most of which contain a Celtic river-name as a specific) most likely arose out of necessity when invading Germanic tribes established new settlements and adopted parts of the local hydronymic and toponymic terminology and attached their own place-name forming elements to them. Latin and French elements number very low in the corpus, and occur more frequently as additional elements, and in such cases their function and status is more like that of an adornment rather than that of an actual place-name formant.

In conclusion, the occurrences of various name-formations in the corpus are summarized in Table 10 on the final page at the end of this chapter.

6.2. Possible directions for future research

In this dissertation, there were a number of tangential topics that were mentioned and that would warrant further research as they could not have been tackled here due to the different focus of the investigation. Below are some further topics for research involving place-names, historical linguistics, sociolinguistics, and code-switching in any combination that I wish to undertake in some form in the future.

In general, from a synchronic aspect, providing a descriptive linguistic account of the process and manifestations of code-switching, i.e. exploring the phonological, morphological, syntactic, and semantic-pragmatic facets of the phenomenon, still offers quite a lot of potential, especially in language pairs and communities hitherto not considered (for instance studying English-Hungarian code-switching among university students in English programs). Similarly, speculations about the theoretical underpinnings as well as the sociolinguistic and psycholinguistic implications of CS are also fruitful areas to investigate. Diachronically speaking, research of code-switching can be expanded to include other levels of language apart from the dominant focus on syntax.

Place-names lend themselves very well to corpus linguistics, however to my knowledge apart from Fekete (2015) no other paper has been published that utilizes corpus linguistics in onomastic research, therefore exploring the topic of corpus onomastics would be a useful direction to take. For instance, aspects of English toponyms which did not form part of the present investigation could be coded in a corpus, even including data points such as the

quality of the soil underneath the settlement or pieces of archeological evidence unearthed, etc. Fictional place-names and what general rules of coinage affect them can also be analyzed quite well from a corpus linguistic aspect.

The parallels and differences between onomastic sound change (as proposed by Clark 1991, Colman 1992, and Coates 2006b) and the obscuration of compound words would also be worth investigating. The exact types of sound change that occur in toponyms could be compared with the types of sound change that affect darkened compounds to see whether or not there are any similarities between the two groups of words. This way the existence of a separate onomastic sound change affecting only proper names could be ascertained and its exact characteristics could be defined. To my knowledge, no research has yet been conducted on the comparison of toponyms with (darkened) compound words. As a side-track the relevance of folk-etymology in the reanalysis of names should also be studied. More generally, based on the present dissertation, a universal description of hybrid names and their emergence could also be provided.

Moving more into the realm of historical linguistics, the etymological background and historical development of specifics and generics used in place-names with special attention to borrowings and doublet formation should also be explored, especially with reference to the relationship between Old English and Old Norse names and name elements.

Finally, the way in which headwords are entered into the source dictionary raises an interesting lexicographic issue in connection with corpus linguistics and digital humanities in general. Many entries state instead of providing a definition “identical in meaning and origin with previous [or sometimes next] name”, and when a source is fed into a corpus in an automated way, such entries can pose serious problems for the tagging and sorting software because they cannot recognize them as valid data points. In the future, digital editions of dictionaries may cope better with problems like this.

In lieu of a snappy and though-provoking witticism to leave the reader with, I would like to quote Baxter’s (2011: 273) more down-to-earth observation in connection with research conducted on the Domesday Book. In order to adequately address all the issues arising from such an endeavor “one would ideally need expertise in Old English, Old Norse, Old French, and medieval Latin, as well as a firm command of the political, social, economic, and legal history of England and her neighbors”. If we extrapolate this to include place-names, then we can also add a solid working knowledge of settlement history, archeology, geology, geography, military history, migration patterns, sociolinguistics, and of course historical linguistics. As it is humanly impossible to be well-versed in all of these areas, I

restricted my investigation to a narrow slice of this gigantic topic. This means that certain boundaries had to be drawn and certain corners had to be cut. The research I conducted in this dissertation was conducted with the optimistic thought and aim that it will hopefully be carried on, refined, and expanded with new perspectives by experts in other fields, so that we can work towards a more complete picture.

	Specific														
	Celtic			Latin			OE			ON			NF		
	pers	app	other ⁶⁰	pers	app	other	pers	app	other	pers	app	other	pers	app	other
Celtic	2	14	136				1	1							
Latin	1	1		2											
OE	6	70	120	1	1		2384	5276	35	214	126	2	8	6	
ON	30	11	8		1		100	152	2	547	446	5	10		
NF	2	1	1				1	4						29	
Other							4	15		19	5				

Table 10. Summary of the various types of place-name formations found in the corpus

⁶⁰ The label “pers” refers to personal names, “app” refers to appellatives, and “other” refers collectively to specifics that are neither personal names nor appellatives (i.e. river-names, names of topographical features, tribal names, etc., with hydronyms being the most frequently occurring type in the “other” category).

VII. Magyar nyelvű összefoglaló

7.1. Kutatói kérdések és a disszertáció célja

A disszertáció célja az Angliában előforduló hibrid helynevek korpuszalapú vizsgálata. A dolgozat újdonságát az adja, hogy egyrészt a helyneveket mint a nyelv szerves részeit kezeli, nem pedig mint a nyelvvel valamiféle speciális viszonyban levő, azon felül álló elemeket. Másrészt korpuszalapú megközelítést alkalmaz a helynevek vizsgálatára, ami a szakirodalomban eddig nem volt jellemző kutatási irányvonal. Harmadrészt a disszertáció fő elméleti keretét a történeti kódváltás területe adja, amelyben ezidáig szinte kizárólag mondattani folyamatokra koncentráló tanulmányok születtek, a dolgozat pedig az összetett szavak szintjén vizsgálja a kódváltás folyamatát. Végezetül pedig az elemzésben a helyneveket elhomályosult összetételekként kezelem, bemutatva fonológiai és morfológiai változásait is.

A vizsgálathoz megfogalmazott kutatói kérdések két csoportra bonthatók: a szakirodalmi és elméleti áttekintés során megválaszolandókra valamint az empirikus korpuszelemzés során megválaszolandókra. Mindkét csoportba három-három kérdés tartozik.

Az elméleti kérdések a következők:

- (1) Hogyan különíthető el egymástól a lexikai kölcsönzés (jövevényszavak) és a történeti kódváltás folyamata? Hogyan építhető be a történeti kódváltás egy szinkron elméleti keretbe? Szükség van-e egyáltalán a két folyamat éles különválasztására?
- (2) Vannak-e olyan szintjei a nyelvnek amelyeket nem érint a kódváltás (fonológia, morfológia)? Hogyan lehet a hibrid helynevekben megfigyelhető kognáta-helyettesítést (ld. Townend 2002) a kódváltás elméleti keretében leírni? Hogyan jelenik meg a kognáta-helyettesítés a hibrid helynevekben?
- (3) Hogyan lehet a hibrid helynevek kialakulását a történeti kódváltás és a történeti szociolingvisztika elméleti keretében elemezni?

Az empirikus kutatáshoz kapcsolódó kérdések pedig a következők:

- (4) Hogyan befolyásolja a kódváltás kimenetelét és a hibrid helynevek keletkezését a hibridizációban részt vevő nyelvek között fennálló rokonsági és szociolingvisztikai kapcsolat?
- (5) Hogyan jelenik meg a történeti kódváltás az angliai hibrid helynevekben és milyen rendszerszerűségek figyelhetők meg?
- (6) Milyen hibridizációs folyamatok figyelhetők meg az angliai helynevek esetében és ezek mit árulnak el a kódváltásról és a nyelvi kontaktusokról amelyek keretében a helynevek keletkeztek?

7.2. A kutatás elméleti háttere

A kutatás elméleti hátterét két fő téma köré lehet csoportosítani: a kódváltással illetve kódkeveredéssel foglalkozó tanulmányok valamint a helynevek és elhomályosult összetételek nyelvi jellemzőit vizsgáló tanulmányok köré. Mivel a rendelkezésre álló szakirodalomban kevés szó esik a hibrid helynevekről, és a fogalom pontos definícióját sem adják meg kellő részletességgel, a disszertációban saját munkadefiníciót fogalmaztam meg. Azon helyneveket tekintem hibrideknek a szó szűk értelmében, amelyekben az egyik tag a másik tagtól eltérő nyelvből származik, és ez az idegen tag nincsen integrálva a másik nyelvbe illetve jövevényszóként sincs jelen az adott nyelv szókincsében. Tágabban vett értelmezésben viszont az integráció hiánya és a jövevényszóként való jelenlét nem kizáró tényezők, tehát olyan helynevek is hibridnek tekinthetők amelyekben az idegen tag már integrálódott a másik nyelvbe. A lényegi különbség, hogy csak a szűk meghatározás engedi meg a kódváltást és kódkeveredést.

Fontos megjegyezni, hogy a kódváltás és a lexikai kölcsönzés rokon folyamatok, amelyeket leginkább az idegen elemek előfordulási gyakorisága tud egymástól elkülöníteni. Ennek értelmében a jövevényszavak gyakran előforduló elemek, amelyek az egynyelvű beszélők szókincsében is megtalálhatók, viszont azok a szavak amelyek kódváltással kerülnek a beszélő megnyilatkozásaiba általában csak kétnyelvű (vagy többnyelvű) beszélők szókincsében találhatók meg. Ezen felül a kódváltott szavak formális szinten nincsenek integrálva a nyelvbe.

A kódváltás („code-switching”) és a kódkeveredés („code-mixing”) fogalmának meghatározására számtalan tanulmány született már, azonban a szakirodalomban még mindig

nincs teljes konszenzus arról, hogy pontosan mit tekintünk kódváltásnak és mit kódkeveredésnek. Az általánosan elfogadott közelítő meghatározás értelmében kódváltásnak az tekintendő, amikor egy adott megnyilatkozáson belül a beszélő két (vagy akár több) különböző nyelv struktúráit használja párhuzamosan, és a váltások mondatok között nagyrészt elkülönülten szerepelnek („intersentential code-switching”, mondatok közötti kódváltás). Ezzel szemben a kódkeveredés alapja ugyanez a folyamat, azonban a váltás nem mondatok között hanem mondatokon belül történik. A hibrid helynevek esetében is kódkeveredésről beszélhetünk, ami nem mondatokon hanem összetett szavakon belül történik.

A kódkeveredés fő elméleti hátterét Muysken (2000) három kategóriája adja, melyek: beszúrás („insertion”), váltakozás („alternation”) és kongruens lexikalizáció („congruent lexicalization”). A beszúrás során egy idegen nyelvből bizonyos elemeket, általában önálló szavakat vagy csoportokat, a másik nyelv szerkezetébe illesztnek be a beszélők a célnyelvi szerkezetekbe integrálva. Váltakozás során nagyobb, magasabb szintű összetevőket (pl. akár egész mondatokat) illesztnek be a beszélők a célnyelvbe, integráció nélkül, általában mondatonként váltakoztatva a nyelveket. Kongruens lexikalizáció pedig általában olyan nyelvek között jön létre, amelyek egymásnak közeli rokonai, gyakran kölcsönös érthetőség is fennáll közöttük valamint szerkezetileg nagyon hasonlítanak egymásra (éppen ez a fajta kódkeveredés gyakran fordul elő dialektusok között is). Ezen folyamat során a két nyelv közös, egymással átfedésben levő szerkezetébe bármelyik nyelvből bekerülhetnek elemek a beszélők megnyilatkozásaiban. Ez a fajta kódkeveredés mondaton belül vagy szavakon belül történik. A disszertációban vizsgált esetekben az angol-skandináv hibrid helynevek nagy valószínűséggel kongruens lexikalizáció során keletkeztek, míg a többi típusú hibrid beszúrás útján jöhetett létre.

A kódváltás elemzéséhez segítségül hívtam még Myers-Scotton (1983, 1989) jelöltség modelljét („markedness theory”) és a mátrix nyelv (1993) elméletét. A jelöltség modell értelmében minden kétnyelvű megnyilatkozás során a beszélgetőpartnerek identitásukat ütköztetik egymással az által hogy az adott nyelvi közösség kétnyelvűsége vonatkozó normáinak meg akarnak-e felelni vagy nem. Ez azt jelenti, hogy ha a közösségben elfogadott és elterjedt a kétnyelvűség és a kódváltás, akkor ha a beszélő kódváltást használ egy megnyilatkozásában akkor azzal nem szegi meg a nyelvi normát (tehát megnyilatkozása jelöletlen [„unmarked”] lesz) és nem sérül a pragmatikai értelemben vett homlokzata. Ugyanakkor ha egy közösségben nem elterjedt és nem normatív a kétnyelvűség akkor egy kódváltást tartalmazó megnyilatkozás jelöltté („marked”) válik, a beszélgetőpartnerek közötti viszony egyensúlya felborul. Ezen elmélet úgy függ össze a hibrid helynevekkel, hogy ha egy

közösségekben elterjedt a kétnyelvűség és normatív a kódváltás, akkor az így keletkezett helyneveket is el fogja fogani az adott közösség.

Myers-Scotton (1993) másik elmélete a mátrix nyelv modell pedig azt takarja, hogy minden kétnyelvű megnyilatkozásban van egy domináns nyelv (mátrix nyelv) amely a megnyilatkozás nyelvtani hátterét adja és amelyik a beszélő domináns nyelve, valamint egy beágyazott nyelv („embedded language”) amelyikből az idegen elemeket a mátrix nyelve illeszti a beszélő.

A disszertációban vizsgált hibrid helynevek nyelvészeti hátterének alapja az a feltételezés, hogy ezek a toponímák valójában elhomályosult összetételek, amelyek keletkezésükkor szemantikailag áttetszőek voltak és a beszélők nem pusztán jelentés nélküli címkékként használták őket, hanem fontos leíró szereppel és árnyalatnyi különbségek kifejezésére is képes szemantikai tartalommal bírtak (vö. Hough 2012a). A helynevek fejlődésük során (a többi elhomályosult szóösszetételhez hasonlóan) fokozatosan elveszítik a szemantikai kapcsolatot a bennük megtalálható szavak szabad megfelelőivel aminek következtében ugyanazok a hangváltozások és morfológiai változások zajlanak le bennük mint a nyelv többi szavában, viszont a helyneveket ezek a változások más módon és mértékben érintik, mint a szabad lexikai morfémákat (Clark 1991, Colman 1992, and Coates 2006b).

7.3. A korpusz bemutatása

Az empirikus kutatás alapját képező korpuszt kiegészít Ekwall (1980) klasszikus szótárából, nagyrészt pedig a mára új standarddá vált angol helynevek szótárából (Mills 2011), valamint Gillian Fellows-Jensen (1972, 1978, 1985) angliai skandináv helynevekkel foglalkozó regionális tanulmányaiból állítottam össze. A korpusz alapja egy XML dokumentum (Extensible Markup Language, kiterjeszhető jelölőnyelv) amelyet XSL transzformációk segítségével alakítottam adatokká. Az XML adatbázis legnagyobb előnye a testreszabhatósága, ugyanis a felhasználó által definiált tag-ekkel bármilyen típusú és jellegű információ kódolható. A korpuszban minden helynév esetében a következő adatokat kódoltam: név, első előfordulási alak, első előfordulás, jelentés, első tag („specific”), első tag nyelve, első tag típusa (pl. személynév, köznévi, stb.), második tag („generic”), második tag nyelve, földrajzi régió amelyben a helynév előfordult, egyéb megjegyzések.

A korpuszban szereplő helynevek első 10%-át kézzel kódoltam, majd pedig a fennmaradó neveket nagyrészt automatizálva Mills (2011) szótárának digitális verzióját

felhasználva az Excel 2016 segítségével. Az első 10%-os „próbakorpuszt” azért volt szükséges kézzel bevinni mert ilyen módon lehetőségem volt az automatizálás előtt felmerülő hibákat és hiányosságokat észre venni és javítani. A korpusz összesen 10.311 helynevet tartalmaz Anglia területéről, tehát skót, ír, cornwalli és walesi nevek nem szerepelnek benne. A korpuszt kvantitatív módszerekkel elemeztem, főként előfordulási gyakoriságot és type-token arányokat számoltam. A korpusz összeállításához és elemzéséhez a Microsoft XML Notepad 2007-et, a Notepad++ 7.4.2-es verzióját, a TextPad 8.1.1-es verzióját, valamint a Microsoft Excel 2016-os verzióját használtam fel. A korpusz (és egyben a disszertáció) magját egy korábbi tanulmányom adja (Fekete 2015) melyben kizárólag angol-skandináv hibrid helynevekkel foglalkoztam, szintén korpuszalapú megközelítésben. Ott azonban még nem XML típusú korpuszal dolgoztam, hanem egy Excel alapúval, amit a disszertációhoz készített korpuszba átkonvertáltam.

7.4. Az empirikus kutatás eredményei, konklúziók és kitekintés

A korpuszelemzésből kiderült, hogy az óészaki nyelven kívül az angollal kontaktusba kerülő egyik nyelv (azaz kelta, latin, francia) sem tudott érzékelhető és számottevő hatást kifejteni az angol helynevekre. Ennek oka a szubsztrátum és szupersztrátum hatásokban keresendő, valamint az egymással kapcsolatba kerülő nyelvek szerkezeti hasonlóságaiban. Az óangol nyelv a keltához képest szupersztrátum volt, tehát a kelta nyelvből nagyon kevés jövevényszó érkezett az angolba, azonban a kelta strukturális hatása jelentősnek mondható. Ezzel szemben a franciához képest az angol szubsztrátum volt, tehát a francia strukturális hatása kisebb volt, azonban kimagasló mennyiségű francia jövevényszót vett át az angol. A latin nyelv esetében pedig – Vennemann (2011) terminusával élve – „presztrátum” hatás érvényesült, vagyis a latin annak presztízse okán volt képes jelentős mértékben hatni az angolra ami főleg a szókölcsonzésekben jelenik meg. Annak ellenére, hogy a kelta nyelv szubsztrátum volt az óangolhoz képest, a fent említett három nyelv közül a kelta hatás jelenik meg legnagyobb mértékben az angol helynevekben. Az angol helynevekben előforduló kelta elem az esetek túlnyomó többségében folyónév vagy más természeti képződmény neve, tehát ezek a hibrid nevek úgy keletkeztek, hogy a benyomuló germánok átvették a kelta lakosság folyóneveit, amelyekhez később saját helynévképző elemeiket illesztették. Ezek a nevek is hibrideknek tekinthetők, azonban meg kell jegyezni, hogy a tisztán kelta helynevekkel ellentétben a kelta folyóneveknek szükségszerűen túl kellett élniük a germán hódítást, ugyanis azok a helyi természeti környezet szerves részét képezték.

Latin nevek elenyésző számban fordulnak elő a korpuszban (összesen 11 darab, ezek közül mindössze 3 hibrid), aminek oka, hogy az óangol nyelv beszélői nem voltak közvetlen kontaktusban a latin beszélőivel, az a nyelv csak presztrátumként volt jelen. Ezt az is megerősíti, hogy bár nagyon kevés latin név van a korpuszban, 110 esetben fordul elő, hogy egy már meglévő helynévhez latin kiegészítő elem társul. Ezek az elemek általában az adott település „gazdájára” (pl. *episcopus* „püspök”, *regis* „király”) vagy méretére (pl. *magna* „nagy”) utalnak és céljuk a hasonló nevű települések megkülönböztetése egymástól. Összesen kilencféle latin nyelvű kiegészítő elem fordul elő a korpuszban, melyek főként az uralkodóosztály és az egyház szemantikai mezőjébe tartoznak. A latin presztrátum jellege abban mutatkozik meg, hogy a tipikusan latinhoz köthető (szociolingvisztikai értelemben vett) domáinek szavai jelennek meg kiegészítő elemként, valamint hogy a méret megjelölésére használható angol nyelvű szó helyett azok latin megfelelőjét használták.

A francia nevek szintén alacsony számban fordulnak elő (összesen 57 darab), melynek oka, hogy a normann hódítás (1066) után a teljes angolszász uralkodóosztályt normann franciák és szövetségeseik váltották le, amelyhez azonban nem társult migráció és számottevő letelepedés (ellentétben az óészaki hódítással). Annak ellenére, hogy a normann francia szupersztrátumként funkcionált az angolhoz képest, a helynevekre gyakorolt hatása nagyon csekély.

A legjelentősebb hatást az óészaki nyelv gyakorolta az óangol helynevekre. Ez abból is látható, hogy míg összesen 198 kelta-angol, 3 latin-angol, 17 francia-angol hibrid helynevet tartalmaz a korpusz, addig 596 skandináv-angol hibrid név található az adatbázisban. A jelentős skandináv hatás hátterében az áll, hogy egyrészt az óangol és az óészaki nyelv a kontaktus időtartama alatt nagyban hasonlított egymásra, feltételezhetően kölcsönös érthetőség és valamilyen szintű kétnyelvűség is fennállt a két csoport esetében, másrészt pedig a skandináv hódítás jelentős migrációval, letelepedéssel és az alsóbb társadalmi rétegek szükségszerű asszimilációjával járt. Ezen felül az óészaki nyelv bizonyos mértékben szupersztrátum volt az óangolhoz képest, azonban emellett adsztrátum hatás is megjelenik, tehát a két nyelv nagyrészt hasonló presztízst és státuszt képviselt a közösségben.

A korpuszelemzésből kiderül, hogy az egyes nyelvek egymáshoz képesti szociolingvisztikai és kontaktnyelvészeti viszonya befolyásoló hatással bír a helynevekben megjelenő hatásokra. A szubsztrátumnyelvek (jelen esetben kelta és a kelták előtti nyelvek) korlátozott lexikai hatás fejtenek ki a szupersztrátumaikra, azonban a helynevekben megjelennek folyónevek és más természeti képződmények neveivel. A szupersztrátumnyelv (jelen esetben francia) jelentős lexikai hatást fejt ki a szubsztrátumra, azonban csak

elhanyagolható mértékben jelenik meg a helynevekben. Egyedül az adsztrátum képes jelentős hatást gyakorolni egy adott nyelv helynévkészletére, aminek oka a nyelvek hasonló presztízse és funkciója valamint hogy a másik nyelv beszélői nagy számban telepednek le az adott térségben. Ezen felül a két nyelv hasonlósága is elősegítette a hibrid helynevek kialakulását.

A kódváltás tekintetében megállapítható, hogy a skandináv-angol hibrid helynevek nagy részében beszélhetünk a kódváltás jelenségéről. Ez abban nyilvánul meg, hogy az érintett nevekben megtalálható óészaki helynévképzők közül két elem gyakorisága kimagasló (*bý* „falu” és *þorp* „másodlagos település”) a többi meredeken zuhanó tendenciát mutat, és sok képző csak egyszer fordul elő a korpuszban, ami nagy valószínűséggel kódváltásra és nem pedig lexikai kölcsönzésre utal. Azokban a helynevekben, amelyekben az első tag az óészaki nyelvből származik, a helynévképző pedig óangol ugyanez a tendencia figyelhető meg. A kelta hibrid nevek esetében is hasonló eloszlásról beszélhetünk, tehát nagy számú óangol helynévképző adódik a kelta folyónevekhez. A latin és normann francia nevek alacsony számából kifolyólag nem lehet teljes bizonyossággal megállapítani, hogy kódváltás vagy kölcsönzés történt-e, azonban valószínűsíthető, hogy ezek egyedi, elszigetelt esetek voltak.

Végezetül pedig néhány szó a téma további kutatási lehetőségeiről. A helynevekről, mint említettem, tudomásom szerint még nem készült korpusznyelvészeti elemzés, a téma és a terület jórészt kiaknázatlan. A helynevek kiváló alapot szolgáltatnak korpuszok összeállításához, ugyanis maguk a nyelvi egységek kicsik, viszont rengetegféle adatot lehet hozzájuk kódolni (pl. a település nevének és a név nyelvészeti jellemzőin kívül a méretét, típusát, demográfiai jellemzőit, stb.) ezáltal tágabb kontextusban és tágabb összefüggésekben lehet láttatni a helynevek jellemzőit és akár kialakulásukat is. Ezen felül pedig a (hibrid) helynevek kialakulásának szociolingvisztikai tulajdonságait is érdemes vizsgálni, valamint a nevek fonológiai, morfológiai és szemantikai jellemzői is ígéretes kutatási területek, főként az összetételek elhomályosulására koncentrálva.

References

- Adams, Douglas Q. & Mallory, James P. (eds.) 1997. *Encyclopedia of Indo-European Culture*. Chicago: Fitzroy Dearborn Publishers.
- Algeo, John. 2010. "Is a theory of names possible?" *Names* 58.2 (2010): 90-96.
- Andersson, Thorsten, Eva Brylla & Allan Rostvik (eds.). 1980. *NORNA-rapporter 17, Ortnamn och Språkkontakt: Handlingar från NORNA:s sjätte symposium i Uppsala 5-7 maj 1978*. Uppsala: Almqvist & Wiksell.
- Appel, René & Pieter Muysken. 2005. *Language Contact and Bilingualism*. Amsterdam: Amsterdam University Press.
- Aybes, C. & Derek William Yalden. 1995. "Place-name evidence for the former distribution and status of wolves and beavers in Britain." *Mammal Review*. 25.4 (1995): 201-227.
- Azuma, Shoji. 1996. "Free morpheme constraint revisited." *World Englishes*. 15.3 (1996): 361-367.
- Backus, Ad. 2015. "A usage-based approach to code-switching: The need for reconciling structure and function." In: Stell, Gerald & Kofi Yakpo (eds.). *Code-switching between Structural and Sociolinguistic Perspectives*, 19-37. Berlin: Walter de Gruyter.
- Backus, Ad & Margaret Dorleijn. 2009. "Loan translations versus code-switching." In: Bullock & Toribio 2009: 75-93.
- Bauer, Laurie. 1988. *Introducing Linguistic Morphology*. Edinburgh: Edinburgh University Press.
- Bauer, Laurie. 2008. "Exocentric compounds." *Morphology*. 18 (2008): 51-74.
- Baukó, János. 2015. *Bevezetés a szocioonomasztikába*. Nyitra: Konstantin Filozófus Egyetem.
- Baxter, Stephen. 2011. "The making of Domesday Book and the languages of lordship in conquered England." In: Tyler 2011: 271-308.
- Bába, Barbara. 2013. "Földrajzi köznevek mint lexikális helynévformánsok." *Magyar Nyelvjárások*. 51 (2013): 103-110.
- Bergs, Alexander & Laurel J. Brinton. 2012. *English Historical Linguistics: An International Handbook. Volume 1*. Berlin: Mouton de Gruyter.
- Bertacca, Antonio. 2007. "Contact vs. internal dynamics in the typological shift of English." *Linguistica e Filologia*. 25 (2007): 21-39.
- Bloomfield, Leonard. 1935. *Language*. London: Allen & Unwin.

- Blanár, Vincent. 2009. "Proper names in the light of theoretical onomastics." In: *Namenkundliche Informationen 95/96*. Ed. by Ernst Eicher, Karlheinz Hengst and Dietlind Krüger, 89-159. Leipzig: Leipziger Universitätsverlag.
- Bloomer, Robert K. 1999. "The obscured nominal compounds in German and English: Synchronic stations and processes of change." *Sprachtypologie und Universalienforschung (STUF)*. 52.1 (1999): 52-63.
- Bölcskei, Andrea. 2011. "Törzs- és népnévi helynevek a Brit-szigeteken." *Helynévtörténeti Tanulmányok*. 6 (2011): 197-212.
- Bölcskei, Andrea. 2012. "Helynévi szubsztrátumok a Brit-szigeteken." *Helynévtörténeti Tanulmányok*. 8 (2012): 151-169.
- Bölcskei, Andrea. 2013. "A kelta vallásosság helynévi nyomai a Brit-szigeteken." *Helynévtörténeti Tanulmányok*. 9 (2013): 169-184.
- Brinton, Laurel J. & Elizabeth Closs Traugott. 2005. *Lexicalization and Language Change*. Cambridge: Cambridge University Press.
- Bullock, Barbara E. & Almeida Jacqueline Toribio (eds.). 2009. *The Cambridge Handbook of Linguistic Code-switching*. Cambridge: Cambridge University Press.
- Bullock, Barbara E. & Almeida Jacqueline Toribio. 2009a. "Themes in the study of code-switching." In: Bullock & Toribio 2009: 1-19.
- Cain, Christopher M. 2016. "þæt is on englisc: performing multilingualism in Anglo-Saxon England." In: *Imagining Medieval English: Language Structures and Theories, 500-1500*. Ed. by Tim William Machan, 81-99. Cambridge: Cambridge University Press.
- Cameron, Kenneth (ed.). 1975. *Place-Name Evidence for the Anglo-Saxon Invasion and Scandinavian Settlements: Eight Studies*. Nottingham: English Place-Name Society, 55-98.
- Cameron, Kenneth. 1996. *English Place-Names*. London: Batsford.
- Cannon, John & Anne Hargreaves. 2001. *The Kings & Queens of Britain*. Oxford: Oxford University Press.
- Clanchy, Michael T. 2014. *England and its Rulers: 1066-1307*. Oxford: Wiley-Blackwell.
- Clark, Cecily. 1991. "Towards a reassessment of "Anglo-Norman influence on English place-names." In: Ureland & Broderick 1991: 275-293.
- Clark, Cecily. 1992a. "Onomastics." In: Hogg 1992: 452-489.
- Clark, Cecily. 1992b. "The myth of the "Anglo-Norman scribe." In: Rissanen, Ihalainen, Nevalainen & Taavitsainen 1992: 117-129.

- Clark, Cecily. 1992c. "Domesday Book – A great red herring: Thoughts on some late-eleventh-century orthographies." In: Hicks, Carola (ed.). *England in the Eleventh Century*. Stamford: Paul Watkins Publishing. 317-331.
- Clyne, Michael. 1987. "Constraints on code switching: how universal are they?" *Linguistics*. 25 (1987): 739-764.
- Coates, Richard. 1987. "Pragmatic sources of analogical reformation." *Journal of Linguistics*. 23.2 (1987): 319-340.
- Coates, Richard. 2006a. "Names." In: Hogg & Denison 2006: 312-383.
- Coates, Richard. 2006b. "Afon Ystwyth and onomastic sound change." *Anuario del Seminario de Filología Vasca "Julio de Urquijo"*. 40 (2006): 265-272.
- Coates, Richard. 2007. "Invisible Britons: The view from linguistics." In: Higham 2007, 172-191.
- Colman, Fran. 1992. *Money Talks: Reconstructing Old English*. Berlin: Mouton de Gruyter.
- Colman, Fran. 2012. "Monetary policy and Old English dialects." In: Hegedűs & Fodor 2012: 73-93.
- Colman, Fran. 2014. *The Grammar of Names in Anglo-Saxon England: The Linguistics & Culture of the Old English Onomasticon*. Oxford: Oxford University Press.
- Cox, Barrie. 1980. "Aspects of place-name evidence for early medieval settlement in England." *Viator*. 11 (1980): 35-50.
- Crystal, David. 2011. *A Dictionary of Linguistics and Phonetics*. Sixth Edition. Oxford: Blackwell.
- Čermák, Jan. 2005. "Notes on obscured compounds in Old English." In: *Patterns: A Festschrift for Libuše Dušková*. Ed. by Jan Čermák, Aleš Klégr, Markéta Malá & Pavlína Šaldová, 35-45. Prague: Ústav anglistiky a amerikanistiky FF UK.
- Dalberg, Vibeke. 2008 [1977]. "The psychology of place-name changes." In: *Name and Place: Ten essays on the dynamics of place-names*. Ed. and translated by Gillian Fellows-Jensen, Peder Gammeltoft, Bent Jørgensen and Berit Sandnes, 51-60. Copenhagen: University of Copenhagen.
- Dawson, Hope C. 2003. "Defining the outcome of language contact: Old English and Old Norse." *Ohio State University Working Papers in Linguistics* 57 (2003): 40-57.
- Ditrői, Eszter. 2017. *Helynévrendszerek modellalapú vizsgálata: A helynévminták összevető analízise statisztikai megközelítésben*. Debrecen: Debreceni Egyetemi Kiadó.
- Ekwall, Eilert. 1980. *The Concise Oxford Dictionary of English Place-Names*. Fourth Edition. Oxford: Clarendon Press.

- Fekete, Tamás. 2015. "Anglo-Scandinavian code-mixing in English place-names." *Studia Anglica Posnaniensia*. 50.4 (2015): 5-23.
- Fellows-Jensen, Gillian. 1968. *Scandinavian Personal Names in Lincolnshire and Yorkshire (SPNLY)*. Copenhagen: Akademisk Forlag.
- Fellows-Jensen, Gillian. 1972. *Scandinavian Settlement Names in Yorkshire (SSNY)*. Copenhagen: Akademisk Forlag.
- Fellows-Jensen, Gillian. 1978a. *Scandinavian Settlement Names in the East Midlands (SSNEM)*. Copenhagen: Akademisk Forlag.
- Fellows-Jensen, Gillian. 1978b. "Place-names and settlement in the North Riding of Yorkshire." *Northern History*. 14. 1 (1978): 19-46.
- Fellows-Jensen, Gillian. 1980. "Conquests and the place-names of England, with special reference to the Viking settlements." In: Andersson, Brylla & Rostvik 1980: 192-210.
- Fellows-Jensen, Gillian. 1985. *Scandinavian Settlement Names in the North West (SSNNW)*. Copenhagen: C. A. Reitzels Forlag.
- Fellows-Jensen, Gillian. 1987. "York." *Leeds Studies in English*. 18 (1987): 141-155.
- Fellows-Jensen, Gillian. 1995a. "The light thrown by the early place-names of Southern Scandinavia and England on population movement in the Migration Period." In: *Nordwestgermanisch: Reallexikon der germanischen Altertumskunde*. Ed. by Edith Marhold & Christiane Zimmermann, 57-77. Berlin: de Gruyter.
- Filppula, Markku. 2008. "The Celtic Hypothesis hasn't gone away: New perspectives on old debates." In: *English Historical Linguistics 2006: Selected Papers from the Fourteenth International Conference on English Historical Linguistics (ICEHL 14), Bergamo, 21-25 August 2006. Vol 3: Geo-Historical Variation in English*. Ed. by Marina Dossena, Richard Dury & Maurizio Gotti, 153-170. Amsterdam: John Benjamins.
- Filppula, Markku, Juhani Klemola & Heli Paulasto. 2008. *English and Celtic in Contact*. New York – London: Routledge.
- Gammeltoft, Peder. 2007. "Scandinavian naming-system in the Hebrides – A way of understanding how the Scandinavians were in contact with Gaels and Picts?" In: *West over Sea: Studies in Scandinavian Expansion and Settlement Before 1300*. Ed. by Beverley Ballin Smith, Simon Taylor & Gareth Williams, 479-497. Leiden: Brill.
- Gardner-Chloros, Penelope & Malcolm Edwards. 2004. "Assumptions behind grammatical approaches to code-switching: When the blueprint is a red herring." *Transactions of the Philological Society*. 102.1 (2004): 103-129.
- Gardner-Chloros, Penelope. 2009. *Code-switching*. Cambridge: Cambridge University Press.

- Grice, Herbert Paul. 1975. "Logic and conversation." In: *Syntax and Semantics, Volume 3, Speech Acts*. Ed. by Peter Cole & Jerry L. Morgan, 41-58. New York: Academic Press.
- Hadley, Dawn M. 2000. *The Northern Danelaw: Its Social Structure c.800-1100*. London: Leicester University Press.
- Hadley, Dawn M. 2002. "Viking and native: re-thinking identity in the Danelaw." *Early Medieval Europe*. 11.1 (2002): 45-70.
- Haugen, Einar. 1950. "The analysis of linguistic borrowing." *Language*. 26.2 (1950): 210-231.
- Hegedűs, Irén & Fodor Alexandra (eds.). 2012. *English Historical Linguistics 2010: Selected Papers From The Sixteenth International Conference on English Historical Linguistics (ICEHL 16), Pécs, 23-27 August 2010*. Amsterdam: John Benjamins.
- Hickey, Raymond. 1995. "Early contact and parallels between English and Celtic." *Vienna English Working Papers*. 4.2 (1995): 87-119.
- Hickey, Raymond (ed.). 2010. *The Handbook of Language Contact*. Chichester: Wiley-Blackwell.
- Hickey, Raymond. 2010a. "Language contact: Reconsideration and reassessment." In: Hickey 2010: 1-29.
- Higham, Nick (ed.). 2007. *Britons in Anglo-Saxon England*. Woodbridge: Boydell Press.
- Hoffmann, István. 2007. *Helynevek nyelvi elemzése*. Budapest: Tinta Könyvkiadó.
- Hogg, Richard M. (ed.). 1992. *The Cambridge History of the English Language: Volume I, The Beginnings to 1066*. Cambridge: Cambridge University Press.
- Hogg, Richard M. 2011. *A Grammar of Old English. Volume 1: Phonology*. Oxford: Wiley-Blackwell.
- Hogg, Richard & David Denison (eds.). 2006. *A History of the English Language*. Cambridge: Cambridge University Press.
- Holzer, Georg. 2015. "Mittelalterliche slavisch-deutsche Zweisprachigkeit in Österreich im Lichte onomastischer Mischbildungen." *Вопросы ономастики* 18 (2015): 7-16.
- Hough, Carole. 2000. "Towards and explanation of phonetic differentiation in masculine and feminine personal names." *Journal of Linguistics*. 36 (2000): 1-11.
- Hough, Carole. 2012a. "Celts in Scandinavian Scotland and Anglo-Saxon England: Place-names and language contact reconsidered." In: Stenroos, Mäkinen & Særheim 2012: 3-22.
- Hough, Carole. 2012b. "Onomastics" In: Bergs & Brinton 2012: 212-223.

- Janzén, Assar. 1972. "The Viking colonization of England in the light of place-names." *Names*. 20.1 (1972): 1-25.
- Jones, Michael. 1987. "The failure of Romanization in Celtic Britain." *Proceedings of the Harvard Celtic Colloquium*. 7 (1987): 126-145.
- Jones, Glanville R. J. 1990. "Celts, Saxons and Scandinavians." In: Dodgshon & Butlin 1990: 45-68.
- Kitson, Peter R. 1996. "British and European river-names." *Transactions of the Philological Society*. 94.2 (1996): 73-118.
- Kolehmainen, Leena & Janne Skaffari. 2016. "Multilingual practices in contemporary and historical contexts: interfaces between code-switching and translation." *Multilingua*. 35.2 (2016): 123-135.
- Kunter, Gero. 2011. *Compound Stress in English: The Phonetics and Phonology of Prosodic Prominence*. Berlin: Walter de Gruyter.
- Laker, Stephen. 2010. *British Celtic Influence on English Phonology*. PhD Dissertation, Leiden University.
- Langendonck, Willy van. 2007. *Theory and Typology of Proper Names*. Berlin: Mouton de Gruyter.
- Lipka, Leonhard, Susanne Handl & Wolfgang Falkner. 2004. "Lexicalization & Institutionalization: The state of the art in 2004." *SKASE Journal of Theoretical Linguistics*. 1.1 (2004): 2-19.
- Lipski, John M. 2009. "Fluent dysfluency as congruent lexicalization: a special case of radical code-mixing." *Journal of Language Contact*. 2.2 (2009): 1-39.
- Loyn, Henry. 1994. *The Vikings in Britain*. Oxford: Blackwell Publishers.
- Lutz, Angelika. 2012. "Norse influence on English in the light of general contact linguistics". In: Hegedüs & Fodor 2012: 15-43.
- Lutz, Angelika. 2013. "Language Contact and Prestige." *Anglia*. 131.4 (2013): 562-590.
- Mańczak, Witold. 1985. "Semantic development of borrowings." In: *Historical Semantics – Historical Word-Formation*. Ed. by Jacek Fisiak, 367-375. Berlin: Walter de Gruyter.
- Martynenko, Irina Anatolyevna. 2015. "Hybrid toponyms as a result of the merger of two cultures (based on the Hispanic place names of the USA)." *Russian Linguistic Bulletin*. 1.1 (2015): 12-14.
- Mills, Anthony David. 1998. *The Oxford Dictionary of Place Names*. Second Edition. Oxford: Oxford University Press.

- Mills, Anthony David. 2011. *A Dictionary of British Place-Names*. First Revised Edition. Oxford: Oxford University Press.
- Minkova, Donka & Robert Stockwell. 2009. *English Words: History and Structure*. Second Edition. Cambridge: Cambridge University Press.
- Moskowich, Isabel. 2012. *Language Contact and Vocabulary Enrichment: Scandinavian elements in Middle English*. Frankfurt am Main: Lang.
- Muir, Richard. 1998. "Village evolution in the Yorkshire Dales." *Northern History*. 34.1 (1998): 1-16.
- Mutschmann, Heinrich. 1913. *The Place-names of Nottinghamshire*. Cambridge: Cambridge University Press.
- Muysken, Pieter. 2000. *Bilingual Speech: A typology of code-mixing*. Cambridge: Cambridge University Press.
- Muysken, Pieter. 2011. "Code-switching." In: *The Cambridge Handbook of Sociolinguistics*. Ed. by Rajend Mesthrie, 301-314. Cambridge: Cambridge University Press.
- Myers-Scotton, Carole. 1983. "The negotiation of identities in conversation: a theory of markedness and code choice." *International Journal of the Sociology of Language*. 44 (1983): 115-136.
- Myers-Scotton, Carole. 1989. "Codeswitching with English: types of switching, types of communities." *World Englishes*. 8.3 (1989): 333-346.
- Myers-Scotton, Carole. 1993a. *Duelling Languages: Grammatical Structure in Codeswitching*. Oxford: Clarendon.
- Myers-Scotton, Carole. 1993b. "Common and uncommon ground: Social and structural factors in codeswitching." *Language in Society*. 22 (1993): 475-503.
- Myers-Scotton, Carole. 2006. *Multiple Voices: An Introduction to Bilingualism*. Oxford: Blackwell.
- Nicolaisen, Wilhelm Fritz Hermann, Margaret Gelling & Melville Richards. 1970. *The Names of Towns and Cities in Britain*. London: Batsford.
- Nicolaisen, Wilhelm Fritz Hermann. 1974. "Names as verbal icons." *Names: A Journal of Onomastics*. 22.3 (1974): 104-110.
- Nicolaisen, Wilhelm Fritz Hermann. 1976. "Place-name legends: An onomastic mythology." *Folklore*. 87.2 (1976): 146-159.
- Nicolaisen, Wilhelm Fritz Hermann. 1978. "Are there connotative names?" *Names: A Journal of Onomastics*. 26.1 (1978): 40-47.

- Nicolaisen, Wilhelm Fritz Hermann. 1980. "Onomastic dialects." *American Speech*. 55.1 (1980): 36-45.
- Oftedal, Magne. 1980. "Scandinavian place-names in Celtic territory. An attempt at a linguistic classification." In: Andersson, Brylla & Rostvik 1980: 163-191.
- Parsons, David N. 2011. "On the origins of 'Hiberno-Norse inversion-compounds'." *The Journal of Scottish Name Studies*. 5 (2011): 115-152.
- Pfaff, Carol W. 1979 "Constraints on language mixing: Intrasentential code-switching and borrowing in Spanish/English." *Language*. 55.2 (June 1979): 291-318.
- Poplack, Shana. 1980. "Sometimes I start a sentence in English y termino en Espanol': Towards a typology of code-switching." *Linguistics* 18 (1980): 581-618.
- Poplack, Shana. 2015. "Code switching: Linguistic." In: *International Encyclopedia of the Social & Behavioral Sciences. Second Edition. Volume 3*. Ed. by James D. Wright, 918-925. Amsterdam: Elsevier.
- Poplack, Shana, David Sankoff & Christopher Miller. 1988. "The social correlates and linguistic processes of lexical borrowing and assimilation." *Linguistics*. 26 (1988): 47-104.
- Poussa, Patricia. 1991. "Origins of the non-standard relativizers *what* and *as* in English." In: Ureland & Broderick (1991): 295-315.
- Reaney, Percy Hide. 1987. *The Origin of English Place Names*. Ninth Impression. London, New York: Routledge and Kegan Paul.
- Ringe, Don & Joseph F. Eska. 2013. *Historical Linguistics: Towards a twenty-first century reintegration*. Cambridge: Cambridge University Press.
- Rivet, A. L. F. 1980. "Celtic names and Roman places." *Britannia*. 11 (1980): 1-19.
- Rivet, A. L. F. & Kenneth Jackson. 1970. "The British section of the Antonine Itinerary." *Britannia*. 1 (1970): 34-82.
- Rivet, A. L. F. & Colin Smith. 1982. *The Place-Names of Roman Britain*. London: Batsford.
- Robinson, Orrin W. 1992. *Old English and its Close Relatives*. Stanford: Stanford University Press.
- Roffe, David. 2000. *Domesday: The inquest and the book*. Oxford: Oxford University Press.
- Romaine, Suzanne. 1982. *Socio-historical Linguistics: Its status and methodology*. Cambridge: Cambridge University Press.
- Rot, Sándor. 1992. *Old English*. Budapest: Tankönyvkiadó - Magyar Macmillan.

- Salway, Peter. 2010. "Roman Britain (c. 55 BC – c. AD 440)" In: *The Oxford History of Britain*. Revised Edition. Ed. by Kenneth O. Morgan, 1-59. Oxford: Oxford University Press.
- Sankoff, David & Shana Poplack. 1981. "A formal grammar for code-switching." *Papers in Linguistics: International Journal of Human Communication*. 14.1 (1981): 3-45.
- Sawyer, Peter Hayes. 1971. *The Age of the Vikings*. Second Edition. London: Arnold.
- Schendl, Herbert & Laura Wright. 2011. "Code-switching in early English: Historical background and methodological and theoretical issues." In: *Code-Switching in Early English*. Ed. by Herbert Schendl & Laura Wright, 15-45. Amsterdam: Mouton de Gruyter.
- Schendl, Herbert. 2012. "Literacy, multilingualism and code-switching in early English written texts." In: *Language Mixing and Code-Switching in Writing: Approaches to Mixed-Language Written Discourse*. Ed. by Mark Sebba, Shahrzad Mahootian & Carla Jonsson, 27-43. New York: Routledge.
- Schrijver, Peter. 2002. "The rise and fall of British Latin: Evidence from English and Brittonic." In: *The Celtic Roots of English*. Ed. by Markku Filppula, Juhani Klemola & Heli Pitkänen, 87-110. Joensuu: University of Joensuu.
- Schrijver, Peter. 2007. "What Britons spoke around 400 AD." In: Higham 2007, 165-171.
- Schrijver, Peter. 2014. *Language Contact and the Origins of the Germanic Languages*. New York – London: Routledge.
- Sköld, Tryggve. 1980. "Ortnamn och kodväxling." In: Andersson, Brylla & Rostvik 1980: 266-276.
- Stenroos, Merja, Martti Mäkinen & Inge Særheim (eds.). 2012. *Language Contact and Development around the North Sea*. Amsterdam: John Benjamins.
- Szirmai, Mónika. 2005. *Bevezetés a korpusznyelvészethez*. Budapest: Tinta.
- Thomason, Sarah Grey, and Terence Kaufman. 1991. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
- Thomason, Sarah Grey. 2000. "On the unpredictability of contact effects." *Estudios de Sociolingüística*. 1.1 (2000): 173-182.
- Thomason, Sarah Grey. 2001. *Language Contact*. Edinburgh: Edinburgh University Press.
- Townend, Matthew. 2000. "Viking Age England as a bilingual society." In: Hadley & Richards 2000: 89-105.
- Townend, Matthew. 2002. *Language and History in Viking Age England: Linguistic Relations between Speakers of Old Norse and Old English*. Turnhout: Brepols.

- Townend, Matthew. 2014. *Viking Age Yorkshire*. Pickering: Blackthorn Press.
- Tristram, Hildegard. 2007. "Why don't the English speak Welsh?" In: Higham 2007: 192-214.
- Udolph, Jürgen. 2012. "The colonisation of England by Germanic tribes on the basis of place-names." In: Stenroos, Mäkinen & Særheim 2012: 23-52.
- Ureland, Per Sture & George Broderick (eds.). 1991. *Language Contact in the British Isles: Proceedings of the Eighth International Symposium on Language Contact in Europe, Douglas, Isle of Man, 1988*. Tübingen: Max Niemeyer Verlag.
- Vennemann, Theo. 1994. "Linguistic reconstruction in the context of European prehistory." *Transactions of the Philological Society*. 92.2 (1994): 215-284.
- Vennemann, Theo. 2011. "English as a contact language: typology and comparison." *Anglia*. 129. (2011): 217-257.
- Walther, Hans. 1980. "Zur Problematik, Typologie und Terminologie der sogenannten "Mischnamen" (onymischen Hybride)." In: Andersson, Brylle & Rostvik 1980: 143-162.
- Weinreich, Uriel. 1953. *Languages in Contact: Findings and Problems*. The Hague: Mouton.
- Werner, Otmar. 1991. "The incorporation of Old Norse pronouns into Middle English: Suppletion by loan." In: Ureland & Broderick (1991): 369-401.