Introduction

The availability of corpora has had a profound effect on linguistic methods over the past 30 years. It is probably no exaggeration to say that for historical linguistics, a field in which researchers cannot rely on experimental approaches or introspection, corpora have become the most central way in which language data are accessed. The number and range of historical corpora have seen a tremendous increase over the years. As far as Early Modern English is concerned, the *Corpus Resource Database* hosted by VARIENG lists no fewer than 18 linguistic corpora with Early Modern English data, not including Google Books and *Early English Books Online* (https://varieng.helsinki.fi/CoRD/index.html). They cover diachronic and multi-genre corpora, like the *Helsinki Corpus* (HC) and *A Representative Corpus of Historical English Registers* (ARCHER) as well as specialised thematic corpora, like the *Corpus of Historical English Law Reports* (CHELAR) and the *Quaker Historical Corpus* (QHC).

For the field of historical pragmatics the wide range of domains, text types, genres, and activity types that are represented in these corpora provide ample opportunities to investigate pragmatic phenomena across different contexts. However, applying corpus methods to the study of pragmatics is not without challenges. While corpora make it possible to search for forms, pragmatic functions cannot be retrieved automatically. In addition, corpora tend to present a de-contextualised view of data, which privileges the 'vertical' reading of data across different texts over the 'horizontal' reading of linguistic phenomena in context (Rühlemann and Aijmer, 2014, 3). These by now well-known challenges have been explored within the growing field of corpus pragmatics, both for research on present-day data and historical data (e.g. Aijmer and Rühlemann, 2014; Aijmer, 2018; Andersen, 2011; Jucker, 2013; Landert et al., 2023; O'Keeffe, 2018; Romero-Trillo, 2008; Rühlemann, 2010, 2011, 2019; Suhr and Taavitsainen, 2012; Taavitsainen et al., 2014; Taavitsainen, 2018).

Despite the ever-growing body of research, there still remain many questions and issues to be solved in historical corpus pragmatics. One of these questions is how we can reconcile the need for detailed qualitative analysis that is inherent in pragmatic research with the increasing size of the available data. It is clear, for instance, that new historical corpora include many new, revealing examples of language use in earlier periods, but it is far from clear how such instances can be retrieved, especially if we are interested in realisations of pragmatic functions with forms that have not yet been described systematically. In addition, if detailed manual analysis of text passages is needed for the study of pragmatic functions, it would be desirable to have ways that make it possible to automatically identify text passages that are likely to be especially relevant for the analysis. Pragmatic functions tend to be distributed very unevenly across texts, which means that the study of a random selection or a stratified sample of text passages is often a very inefficient approach. In contrast, basing the text selection on specific contexts (e.g. a certain genre) for which a high presence of a certain pragmatic function is expected results in insights that are restricted to this context and that may confirm preconceived hypotheses, while uses in contexts that do not correspond to our expectations tend to be missed. At the moment, the existing methods provide only limited options for improving the qualitative analysis of corpus data and, thus, developing new methods is highly desirable.

In this book, I apply one such method, which has so far only been discussed briefly (Landert, 2019; Landert et al., 2023). The method takes as its starting point the observation that many pragmatic features tend to occur in clusters. By identifying such clusters of features, it is possible to identify passages that include rich material for qualitative analysis. The analysis of such high-density passages can then lead to new insight into the use of pragmatic phenomena and form the basis for additional corpus-based studies. Thus, the aim is to complement existing methods and to provide an additional, bottom-up and corpus-driven perspective, which makes it possible to combine quantitative corpus pragmatics with qualitative analysis and philological approaches. One of the advantages of the method is that it is scalable: despite the fact that it facilitates detailed manual analysis, the amount of work that has to be invested in the analysis by the researcher does not increase linearly with corpus size. This means that more and larger corpora can lead to richer insight also for studies that rely on detailed qualitative analysis. Methods supporting qualitative analysis are very valuable for the field of historical corpus pragmatics,

which tends to rely on context-dependent interpretations of linguistic examples. Several of the chapters in this book will be based on this method, and the analyses resulting from it will demonstrate that it is a fruitful approach with the potential to complement existing methods in corpus pragmatics.

The pragmatic phenomenon that serves as a test case for the method is epistemic stance, that is, the expression of attitudes concerning the degree of certainty and reliability of information. I investigate epistemic stance in Early Modern English, where we can find passages in which several such expressions cluster together. This results in texts that, from a present-day perspective, may appear as expressing epistemic stance in a remarkably explicit, elaborate, or perhaps even somewhat redundant way. Examples 1.1 to 1.4 present some instances of such passages.

(1.1)   *Trulie I doe verily thinke* that I shall not goe out of my chamber this long time: *perhaps* not at all, which is *more likely,* being troubled with a burning feaver: (PCEEC, Oxinde 109, 1640)

(1.2)   *I know full well,* it will be obiected against me, as it hath beene against others that were men of great learning and iudgement, that my learning and knowledge heerin, is not to be compared with those men, which hold the contrarie opinion, *I grant it to be true,* and I would indeede hold their opinions to be good also, but that *daily experience hath taught me to the contrarie.* (EMEMT, William Clowes, 'Lves venerea', 1596)

(1.3)   But since this is *as evident, as that the Sun shines at Noon-Day*, since *none in this Age deny it, except Persons of lewd Morals, and a reprobate Sense, I think,* the Matter too notorious to be at this Time insisted on, and only desire such, who doubt the Truth hereof, that they would seriously peruse the View of the Immorality and Profaneness of the English Stage, with the Sense of Antiquity on this Argument. (Lampeter, Religion, 'A Sermon Preached in the Parish-Church of St. Butolph's Aldgate …', 1730)

(1.4)   L. Pres. By whom, who did the Dialogue run between? Did Mr. Love do it?
Mr. Adams. *Truely, for my part, I did not then take such speciall notice, that I dare at this time upon my oath, deliver any thing positively* against any particular man, but generally all spake something, *as I remember.*
Att. Gen. Was there any debate to mend the instructions?
Adams. *Truly,* Sir, not *as I remember.*

> Att. Gen. Did not Mr. Gibbons bring the rough draft of the
> Commission?
> Mr. Adams. *Truely, I think* he did, *as I remember* he did, I have
> spoken of the Comission and instructions that were debated then
> and there, and that many of these persons here mentioned, were
> there; but that some of them mentioned, *might* not be there then, *is
> very likely*.
> (CED, Trial, 'Triall of Mr Love', 1651/1652)

The author of the letter in Example 1.1, James Oxinden II, uses a number
of different stance markers when telling his brother how certain he is that
he will not be leaving his chamber for a long time: *Trulie, I doe verily
thinke, perhaps* and *more likely*. The author of the medical treatise quoted
in Example 1.2 uses the lexico-grammatical stance marker *I know (full
well)* and the phrase *I grant it to be true* to express his epistemic stance.
In addition, he refers to his *daily experience* as support of his position,
which he presents as being in opposition to positions held by *men of great
learning and iudgement*. Similarly, the author of the religious pamphlet in
Example 1.3 combines lexico-grammatical stance markers, like *evident* and *I
think* with additional rhetorical strategies, in this case a simile (*as evident, as
that the Sun shines at Noon-Day*) and attacks against anyone who does not
share his position (*none in this Age deny it, except Persons of lewd Morals,
and a reprobate Sense*). While such rhetorical strategies are not stance
markers in the narrow sense, they reinforce the epistemic stance expressed
by the author, which, in this case, is one of great certainty. In contrast,
Example 1.4, which comes from a transcript of a courtroom trial, includes
a speaker who uses stance markers to establish carefully the limited degree
of certainty to which he is willing to commit himself with his statements.

 Passages like those quoted in Examples 1.1 to 1.4 pose a striking contrast
to some of the claims found in previous literature on the diachronic
development of stance marking. For instance, in his quantitative study of
modals and lexico-grammatical stance markers in the ARCHER corpus,
Biber (2004a) finds that except for modal verbs, stance markers have
increased over time. He concludes:

> These developments indicate a general shift in cultural norms: speakers and
> writers are simply more willing to express stance in recent periods than in
> earlier historical periods. (Biber, 2004a, 129–30)

Of course, it is possible that the passages with very explicit stance marking
presented above are not representative for how stance is expressed in Early
Modern English overall. A general low 'willingness' to express stance does

not rule out the existence of some texts and authors that show opposing tendencies. Still, there is reason to argue that the issue of how explicitly stance is expressed in Early Modern English warrants further attention.

One crucial issue that needs to be addressed is the range of expressions that are included in studies of epistemic stance. While Biber (2004a, 109) claims to examine the 'entire system of stance devices', it is clear that his study would only cover a small number of all stance markers used in Examples 1.1 to 1.4. His analysis includes a closed set of easily retrievable lexico-grammatical patterns that are based on previous research of Present-day English. As a consequence, he does not include *truly* (used in Example 1.4) and *verily* (used in Example 1.1), two adverbs that are commonly used as stance markers in Early Modern English, but that have become less common in Present-day English. The adverb *positively* in Example 1.4 is another one not covered by Biber. Since his analysis of stance verbs, adjectives, and nouns used with complement clauses is restricted to instances with overt complementisers (*that*, *to*), he would miss the two instances of the stance marker *I think* in Examples 1.3 and 1.4, which are used with zero complementiser. The same applies to *I know full well* in Example 1.2. Attributive and predicative uses of stance adjectives are not included by Biber either, which excludes *likely* and *true* in Examples 1.1 and 1.2. Likewise, the stance phrases *for my part* and *(not) as I remember* in Example 1.4 would be missed, too. Moreover, rhetorical strategies are not covered at all by Biber (2004a). In sum, his approach would cover at most nine instances of epistemic stance markers in the four passages: *I think*, *shall*, and *perhaps* in Example 1.1, *will*, *would*, and *indeed* in Example 1.2, *would* and possibly *evident* (assuming that the following *that* would be treated as a complementiser of *evident*) in Example 1.3 and *might* as the only stance marker in Example 1.4.

Expanding the inventory of lexico-grammatical constructions might appear to be an obvious solution to addressing this issue. Indeed, Biber's inventory has been expanded in subsequent studies (e.g. Gray et al., 2011; Gray and Biber, 2014), although the expanded inventory has not so far been applied to tracing the diachronic development of stance markers from Early Modern English to Present-day English. However, as I will argue in this book, compiling an inventory of stance markers with the aim of achieving a good quantitative approximation to how stance is expressed is fraught with many problems. The most fundamental of these is the fact that stance expressions in English are context-dependent. In other words, expressions need to be analysed in context when one wants to decide whether or not they are used to mark stance. The frequency with which

given expressions are used to mark stance varies considerably; while some expressions mark stance in the large majority of instances, others do so only rarely. The number of expressions that are rarely or occasionally used as stance markers is considerable and excluding such instances will invariably lead to a distorted picture of how stance is expressed. All this means that reliable quantitative assessments of how frequently stance is expressed are very hard to achieve.

On the level of stance research, the aim of this book is to present new insight into the relation between epistemic stance expressions and the context in which they are used. Throughout the various chapters of this book, I will argue for a context-dependent perspective on stance, and I will conduct a number of different studies that demonstrate the extent to which context interacts with lexical items and lexico-grammatical patterns in stance expressions. This includes every level of context, from the micro-level of co-occurring expressions, to the content and function of a text, its genre, text type, and communicative form, as well as the broader social, cultural, historical, and situational context, and the biographical background of speakers and authors. I will argue that only by taking such contextual factors into account can we truly understand how epistemic stance is expressed and what further functions stance expressions serve.

There are several reasons for basing this research on Early Modern English. Compared to Present-day English, we still know far less about how stance was expressed in earlier periods. Most of the studies that take a diachronic approach have focused on the development of a small set of stance markers, usually markers that still play an important role in Present-day English, such as *I think*, *it seems*, and modal verbs (e.g. Aijmer, 1997, 2009; López-Couso and Méndez-Naya, 2014; Palander-Collin, 1999a, b; Whitt, 2015). Studies that take a comprehensive perspective regarding stance marking in historical texts usually have been restricted to one specific domain or genre, such as medical writing (Alonso-Almeida and Mele-Marrero, 2014; Gray et al., 2011; Taavitsainen, 2000, 2001, 2009), witness depositions (Grund, 2012, 2013, 2021) and personal correspondence (Fitzmaurice, 2003, 2004). As a consequence, we know little overall about how Early Modern English stance expressions depend upon factors such as the genre and the domain in which they are used.

Moreover, Early Modern English presents excellent data for the corpus pragmatic study of context-dependent stance expressions. In Present-day English, epistemic stance is most commonly expressed through modal verbs and lexico-grammatical expressions, such as *I think*, which have

undergone grammaticalisation. In Early Modern English, epistemic uses of modal verbs are less common and the grammaticalisation process of stance expressions such as *I think* is less advanced. As a consequence, compared to Present-day English, we can expect that there is more lexical and structural variation of stance markers in Early Modern English and that context-dependent expressions play a more important role. While this is likely to apply to earlier periods of English, too, Early Modern English has the advantage of offering much larger amounts of data from a broad variety of contexts (see Brinton, 2012, 105). For the research presented in this book, I work with several existing Early Modern English corpora, and I discuss the challenges and demonstrate the value of combining these corpora.

Two theoretical perspectives on stance will be developed as part of this analysis, which, like the method, can be applied beyond Early Modern English. On the one hand, I will provide a reconceptualisation of the concept of certainty that relates the default stance with respect to certainty to contextual factors of the genre. This helps explain not only the frequency with which certainty expressions are found in different genres but also the functions with which they are used. Not least, the reconceptualisation resolves the apparent certainty paradox, that is to say that utterances including certainty markers tend to be perceived as expressing less certainty than unmarked utterances.

The second theoretical perspective concerns a distinction between explicit and implied stance marking. I will argue that certain groups of expressions – evidential expressions, predictions, and some meta-communicative expressions – can be used as indirect expressions of epistemic stance in a systematic way. On the one hand, introducing a distinction between explicit and implied stance marking is helpful for deciding which stance markers to include in a given analysis. On the other hand, the model helps explain the relation between evidential, predictive and meta-communicative expressions and epistemic stance.

As an additional outcome, I will present inventories of epistemic stance markers used in Early Modern English, including both collections of markers that have been studied in previous research, as well as markers I identified in the course of the research presented in this book. This includes a list of 33 lexical items that I identified through my method of retrieving passages with a high density of stance markers. All of these items were used to express epistemic stance in my data, although they have previously been excluded from quantitative studies of stance markers in Early Modern

English. This list, together with other collections of stance markers, can be found in Appendix B.

This book starts with Part I, consisting of two chapters that cover different aspects of relevant theoretical background. Chapter 2 presents a more detailed exploration and discussion of the challenges that the field of historical corpus pragmatics is facing, which relates the aims of this book to the wider research context. The chapter also includes an overview of the methods that have been applied to the study of stance in historical data so far and the presentation of the corpora on which I base most of my analyses. In Chapter 3, I introduce the most important theoretical concepts that play a role in the study of stance, including the relation between epistemicity and evidentiality, the different structural types of lexico-grammatical stance markers and the role of modal verbs in epistemic stance in Early Modern English.

Part I is followed by Part II, which includes two chapters on methodological aspects. Chapter 4 is devoted to the topic of inventories of stance markers and the question of how stance markers can be identified that have not been discussed in the research literature. I discuss four different approaches to the identification of such unexplored stance markers, providing illustrations and sample analyses of each of them. In addition, I introduce a number of classification criteria that can help decide which items to include in a stance inventory, depending on the goals of the study. For, as I will argue in more detail there, compiling a universal and comprehensive stance inventory is not feasible. Instead, inventories should be compiled on a case-to-case basis, including those markers that are most productive for the analysis in question. The chapter concludes with one such inventory, consisting of only a small set of frequent and reliable markers of epistemic stance in Early Modern English, which I compiled for central analyses presented in this book. One of the main contributions of this book is the method of retrieving passages with a high density of stance markers from corpora. This method is discussed and evaluated in Chapter 5.

Part III of the book includes four empirical studies of epistemic stance in Early Modern English. In Chapter 6, I present the results of a first application of the retrieval of high-density passages. In this study, I manually analyse the 42 passages I retrieved from the corpora in order to identify all epistemic stance markers that are used in them, resulting in 33 lexical stance markers that have been excluded from previous stance research. In Chapter 7, I continue the analysis of the high-density passages. This time, the focus lies on contextual factors and, especially, on rhetorical

strategies that tend to co-occur with epistemic stance markers. The chapter reveals differences in the use of stance markers across the different corpora and across the sections within most corpora, and it identifies additional contextual factors that can influence the use of stance markers, including, for instance, cultural shifts in scientific thought-styles and the biographical background of language users. In combination, Chapters 5 to 7 demonstrate how new corpus methods can help exploit the available corpus data, even if corpora are potentially large and even if the pragmatic feature under investigation relies on a great amount of manual analysis.

The remaining two empirical chapters help complete the picture of epistemic stance in Early Modern English with two studies that rely on established corpus-based methods of pragmatic analysis. Chapter 8 is devoted to two meta-communicative markers, *I say* and *I tell (you)*. I excluded meta-communicative expressions from the set of frequent and reliable stance markers, leading to relatively few observations of the role of meta-communicative expressions in stance marking in the high-density passages. In this chapter, I show that meta-communicative expressions such as *I say* and *I tell (you)* are often used as stance markers. This chapter also initiates the empirical study of context in relation to stance, including three different types of contextual factors: the micro-context in the form of modal verbs modifying meta-communicative expressions, the genre context in the form of differences between trial proceedings and comedy plays, and the situational context referring to different speaker roles and activity types. As I will show, all of these influence the way in which *I say* and *I tell (you)* are used to mark stance.

Chapter 9, finally, presents a discussion of certainty markers in Early Modern English. Starting with theoretical considerations concerning the apparent paradox of certainty markers, I explore the notion of a default stance of certainty that is inherent in different contexts, for example in the form of default stances for given genres. The empirical part of the chapter is devoted to an analysis of certainty markers in three genres in which certainty is not taken for granted by default, namely political pamphlets, medical treatises, and trial proceedings. This study presents a final perspective – as far as this book is concerned – on the many ways in which epistemic stance depends on context.