Manfred Markus, Yoko Iyeiri, Reinhard Heuberger & Emil Chamson (eds.). *Middle and Modern English Corpus Linguistics: A Multi-dimensional Approach*. Studies in Corpus Linguistics 50. Amsterdam/Philadelphia: Benjamins, 2012, viii + 287 pp., € 99.00/$ 149.00.

This book unites 17 papers read at a conference with the same title organized by the editors in Innsbruck in 2009. The articles are thematically arranged in four parts and preceded by an excellent introduction, in which the editors outline the structure of the book and sketch the topics of the individual contributions. The bulk of the articles of parts II and III deals with the development and varia-tion of phonological, grammatical and lexical elements in the history of English

with a focus on the periods Middle English (ME), Early Modern English (EModE) and Late Modern English (LModE). The frame parts are about the potential of existing and desirable features of future corpora (part I) and the use of corpora in dialect studies (part IV).

Joan Beal (13–29) argues that the compilation and exploitation of electronic corpora provided insights which would not have been possible in the pre-corpus age. She illustrates her claim with examples from syntax, e.g. the rise of the passive progressive, the demise of multiple negation, and the stigmatization of preposition stranding. She also deplores the lack of corpora which would allow to study features of LModE phonology. As a remedy she envisages the compilation of a corpus derived from pronouncing dictionaries of the 18th century. Yet her claim that such a corpus could reveal details about regional and diachronic variation of 18th century phonology should be treated with caution. As demonstrated before in the case of the decline of multiple negation, it could turn out that in the field of phonology, too, a "prescriptive" corpus need not reflect actual practice.

Stefan Diemer (31–45) complains that spelling variation cannot be studied in the existing corpora of ME. In his view they should contain features like abbreviations, cancellations, font type and size, line and letter spacing, ornamentation and illumination in the form of tags or links to manuscript images. From these features much useful information can be derived about the date of the manuscript, the scribe, and production circumstances. Elaborate initials allow conclusions about the illuminator and hence the scriptorium and the amount of care and time spent on the production of the manuscript. Different degrees of letter spacing help to distinguish between prefixed verbs and particle-verb combinations. By these and other examples the author demonstrates the benefits to be gained from this new type of corpus.

In his contribution Isao Hashimoto (49–57) traces the history of English compound numerals of the type *twenty-one*. Their typical form in Old English (OE) was *one-and-twenty* (type 1), their typical ME and EModE form was *twenty-and-one* (type 2), and from EModE onwards the modern form *twenty-one* (type 3) became the rule. The author takes his data from several ME and EModE Bibles and from their sources, the Hebrew Bible and the Vulgate. His analysis of the Hebrew Bible yields 123 tokens of type 1 and 252 tokens of type 2. The EModE Bibles which are translated from the Hebrew original show the same types, yet with varying shares of type 1 and type 2. It is therefore plausible to argue – as the author does – that the replacement of type 1 by type 2 was triggered by these Bible translations. His argument about the spread of type 3 is more complex. It is based on those ME and EModE Bibles which were translated from the Vulgate. The Vulgate contains two tokens of type 1, 42 tokens of type 2, and

331 tokens of type 3. Here chronology plays a role. The ME Bibles and the Coverdale Bible (1535) translate these numerals by a majority of type 1 and a minority of type 2 tokens, whereas the later Douay-Reims-Bible (1582) has a slight majority of type 3 tokens, a considerable minority of type 1 tokens, and just 20 type 2 tokens. The author takes this distribution as a proof of his claim that type 3 was introduced through Bible translations from the Vulgate. Could type 3 not equally well be explained as a shortened form of type 2 as illustrated in the example *.xxxv. yere* (55) from Tyndale's *Pentateuch* (1530), which was translated from a Hebrew original?

Yoko Iyeiri (59–73) explores the distribution of *that*-clauses, *(for) to*-infinitives and bare infinitives as complements of causative *make* in prose texts of the 15th century. She claims that the replacement of *that*-clauses by *(for) to*-infinitives and of these by bare infinitives follows the same principle: "the binding force between the matrix verb *make* and its complement becomes stronger" (61). Following previous studies which identified the distance between *make* and its complements, the cognitive complexity of the environment, and the co-ordination of complements as factors governing the choice of the complement types, she explores if Rohdenburg's complexity principle also holds for her data. She tests several patterns, and most of them confirm the validity of the complexity principle. Only complements containing the high-frequency verbs *be* and *come* prefer *(for) to*-infinitives, whereas *die* occurs more often with the bare infinitive. This distribution holds regardless of the cognitive complexity of the environment. The author very cautiously speculates if the bare infinitive expressing a close link in complements with *die* may correlate with the causee's inability to escape the causer's decision.

Tine Defour (75–92) traces the semantic-pragmatic development of the adverbs *verily*, *truly*, and *really*. All three forms belong to the semantic field of truth, and in Present-Day English they are used as emphasizing subjuncts in medial position and as disjuncts in initial position. Although they occur more frequently in the written part of the *British National Corpus*, the author bases her diachronic study on the multi-genre *Helsinki Corpus* and on the *Corpus of English Dialogues 1560–1760*. This choice of corpora weakens some of her arguments. Her data allow the conclusion that *verily* developed from a manner adverb to an attitudinal disjunct, but it is unfortunate that the latter use is attested more frequently in EModE than in LModE. This may be due to the fact that her LModE texts are in their majority from the spoken end of the spoken-written scale. Another reason is probably that the dialogue corpus contains only few, if any, religious texts, in which *verily* is particularly frequent. *Truly* is attested as a manner adverb already in OE, but truth-related meanings developed only in ME. The subsequent increasing frequency of *truly* as an attitudinal disjunct cor-

relates with a change of the class of verbs modified by *truly* as a manner adverb. Activity verbs, the preferred verb class in ME, give way to communication verbs in EModE. The decreasing frequency of *truly* in LModE is probably due to the same factors as those mentioned in the context of *verily*. The development of *really* differs from that of the other two forms. In the author's corpus it is not attested before the EModE period, and here it is used as an emphasizer. Already in the second half of the 17th century *really* takes on more pragmatic meanings as an attitudinal disjunct.

Sylwester Łodej (93–108) uses dictionary evidence to show that the words *pope*, *bishop*, and *priest*, which originally denoted church officers (not church offices, as wrongly claimed on p. 94), developed pejorative connotations and furthermore extended their meanings to the secular sphere. These changes are particularly pronounced in two periods, namely in the 16th/17th centuries and in the 19th century. These periods also witnessed the first attestation of a particularly big number of derivatives from the terms under investigation. The author claims that these parallel developments were triggered by historical and political developments associated with the Reformation and the Civil War in the first, and with "the Great Awakening and new Anglo-Catholic movements" (94) in the second period. This claim is plausible for the first period with its decline of the dominance of the clergy and the loss of its reverence by the general public. It is less plausible for the second period with its reorientation towards religion and Christian values. In the second part of his contribution the author tests the reliability of the dictionary information on a corpus of 1,045 comedies first performed between 1500 and 1900. He finds that the peak frequency of his search items in the 17th century is perfectly in line with the first attestations in the *OED*, and that the biggest share of pejorative meanings in the same period supports his claim of the influence of the decreasing dominance of the clergy on the meaning changes of the words under investigation. Unfortunately he does not comment on their decreasing frequencies in the plays first performed in the 19th century.

For his study on ME *anger* and *tēne*, Hans-Jürgen Diller (109–124) collected his evidence from the *Middle English Dictionary* (*MED*) and the *Innsbruck Prose Corpus* of the *Innsbruck Computer Archive of Machine-Readable English Texts* (*ICAMET*). From the senses listed in the *MED* the author infers a large semantic overlap of the terms. From its examples he also infers that after 1400 *tēne* is confined to poetry. The last attestation of *tēne* is from the beginning of the 17th century. With the help of his corpus he succeeds in establishing the following factors which governed the distribution of *anger* and *tēne* and their derivatives in ME: *anger* was preferred in prose, *tēne* in poetry; after 1300 *anger* was very productive in the derivation of adjectives which describe the emotional state of

the experiencer, whereas the few occurrences of the adjectival derivatives of *tēne* describe the causer rather than the experiencer; the adjective *angry* modifies human nouns, the adjectives *tēne* and *tēneful* modify nouns denoting what can be experienced by humans; *tēne* is twice as frequent in object function as *anger*. The author concludes that whereas the morphosyntactic differences between *tēne* and *anger* are compatible with the interpretation of modern psychology of anger as an active, and fear or sadness as a passive emotion, the semantic overlap between *tēne* and *anger* presents a problem. This apparent discrepancy is reconciled by the fuzziness of the concepts of ANGER and SADNESS in modern English.

Namiko Kikusawa (127–139) looks at the distribution of the subjunctive and modal expressions in clauses introduced by *lest*. Although she includes complement clauses and adverbial clauses in her study, *lest* turns out to be a very low frequency element with only 237 occurrences in 116 out of the 129 texts of the *Middle English Prose Corpus* of *ICAMET*. It would be interesting to know why the remaining 13 texts were not included in her corpus. The small number of examples is also surprising when compared to Auer's (2008) figure of adverbial *lest*-clauses in the period 1570–1640 in the *Helsinki Corpus*. The author supports Moessner's (2005) finding in ME conditional clauses that even in late ME the subjunctive of lexical verbs was nearly as frequent as the subjunctive of *be*. It would have been helpful to distinguish here between complement clauses and adverbial clauses, since in the former the decline of subjunctives of lexical verbs sets in earlier (Moessner 2007: 223). A difference between the two clause types is noted by the author in the ratio *subjunctive* vs *modal* expression. Adverbial clauses prefer the former, complement clauses the latter. The extralinguistic parameters, whose influence on the distribution of subjunctives and modal expressions the author investigates, are the text category and the dichotomy *written* vs *spoken*. She reduces the 31 text categories of the *Middle English Prose Corpus* to 12, which she sometimes calls text types, sometimes genres, and of which she discusses in detail those 6 which are attested by more than 10 examples. The text categories with the highest frequencies are religion and romance. The texts of the first category prefer the subjunctive, those of the second modal expressions. The preservation of the subjunctive in religious texts is convincingly interpreted as a sign of their conservative nature. Less convincing is the author's interpretation of the preference of modal constructions in spoken religious texts as a consequence of the property of the oral medium "requir[ing] the power of modals [...] to send messages to the audience" (137). The usually more progressive nature of spoken texts would have provided a more adequate explanation.

The position of the quantifier *all* in ME is the topic of Tomohiro Yanagi's paper (141–155). Its data come from the *Penn-Helsinki Parsed Corpus of Middle English*. In this corpus nearly all instances of nouns modified by *all* follow the pattern *all*-noun. This sequence holds irrespective of whether the noun functions as subject or as object. When the head of the construction is a pronoun, preposition and postposition of the quantifier are equally frequent when the pronoun functions as subject; when it functions as object, the pattern pronoun-*all* is the rule. Non-adjacent *all*-quantifier constructions are more or less confined to subject function, and in the majority of cases the head is a pronoun. Since with pronominal subjects both positions of the quantifier *all* are equally frequent, the author considers several factors as potentially relevant: dialect, composition date, grammatical person of the pronominal head, and clause type. Whereas dialect has no influence on the preferred word order, the texts of the 15th century show a preference for the pronoun-*all* pattern. The same pattern is also preferred in subordinate clauses and when the head of the construction is a first or second person pronoun.

Hans Sauer (157–175) counted the interjections in Chaucer's *Reeve's Tale*, and he classifies his set of 27 examples in several ways, e.g. according to their phonological and morphological structure, their position in sentences, their semantic and pragmatic functions, and their etymology. The research question behind all these classifications is how representative Chaucer's use of interjections in the *Reeve's Tale* is of Chaucer's other works and of ME in general. Although some more figures of interjections obtained from the *MED*, the *Corpus of Middle English Prose and Verse*, and from serveral other sources are provided, an answer to the research question is not given.

Ursula Lutzky (177–189) studies the discourse markers *why* and *what* in EModE drama. Her data come from the *Corpus of English Dialogues 1560–1760* and from the *Penn-Helsinki Parsed Corpus of Early Modern English*. In accordance with previous studies she finds that in her data *why* is much more likely to serve as discourse marker than *what*. The frequency of *why* increases during the period investigated, whereas that of *what* decreases. Both discourse markers occur predominantly in turn-initial position, and they express surprise. Another feature which is shared by both discourse markers is their collocation with forms of address, where they are used as attention-catching devices. They differ, however, in their preference of sentence types, *what* preferring yes/no-questions, *why* statements. In the latter, *why*, but not *what*, can also be used with contrastive or with conclusive function. The author illustrates her analyses with well-chosen examples and provides detailed and convincing interpretations.

Erik Smitterberg (191–206) investigates the distribution of contractions with *not* in the 19th century in the context of growing colloquialization. His data

come from the speech-based text categories drama, fiction, and trials in *A Corpus of Nineteenth-Century English*. They were chosen because in the written text categories of this corpus *not*-contractions are not attested. The author is very careful in the selection of relevant examples. He tries to recover all spelling variants and unorthodox forms like *canna* for *cannot* in his corpus, and he considers only those instances of verb plus *not* where the contracted and the uncontracted form are possible. His quantitative analysis yields a statistically significant increase of *not*-contractions in the first half of the 19th century as well as a statistically significant higher proportion of *not*-contractions in drama texts than in the other two text categories. These results are interpreted as indicators of a growing colloquialization similar to that claimed for the end of the 20th century in previous studies. The unexpectedly low frequency of *not*-contractions in trials may reflect a discrepancy between what was actually said in the trials and what was taken down and printed afterwards.

Manfred Markus's paper (209–224) is an instance of work in progress. It reports on the project of the *English Dialect Dictionary Online*, which is based on Joseph Wright's *English Dialect Dictionary*. The author describes the complexity of the entries in the original dictionary and the wealth of information they contain. Both aspects – the complexity and the wealth of information – presented challenges for the research team engaged in the production of the online version of the dictionary. The envisaged and partly already realized structure of its entries are outlined and illustrated with screenshots. As the team leader of the project, the author is understandably proud of several dialect studies in which preliminary versions of the new dictionary proved already very helpful, and he sees great potential of the new tool for further studies on issues of synchronic, but also of diachronic dialectology.

Emil Chamson's article (225–240) is a spinn-off of his doctoral dissertation, which investigates the influence of continental West Germanic languages on LModE dialects with the help of the original *English Dialect Dictionary*. In the present paper the author describes in much detail which kind of etymological information is contained in this source and how it compares to that of other dictionaries. He finds that in most cases the *English Dialect Dictionary* provides useful and reliable etymological information, which, however, is often less precise and less detailed than that of the *OED*.

Javier Ruano-García (241–256) points out that Joseph Wright used many different sources for the compilation of his *English Dialect Dictionary*, one of them being White Kennett's glossary to *Parochial Antiquities* (1695). In his paper the author tries to find out how many words in the *English Dialect Dictionary* are illustrated with examples from this source and which use Wright made of the information he found there. This research question could only be approached

with the help of the online version of the *English Dialect Dictionary*. Its search tool revealed that the meanings of 89 words are illustrated with examples from Kennett's glossary. The remarkable point behind this figure is that Wright used only citations for words which were marked explicitly as regional and that he faithfully took over Kennett's regional classification. In many cases Kennett's glossary was even the only evidence Wright found for the use of particular words in particular regions.

Clive Upton (257–268) claims that a geographical region cannot be tidily divided into a fixed number of regional dialect areas, and he illustrates his claim with the example of the English Midlands. This is the area where previous studies carried out with traditional methods of dialectology established several isoglosses allegedly mapping the dividing-line between northern and southern dialects. Taken together they form a band of about 120 miles width stretching across the Midlands. This fuzziness of the phonological evidence and additionally some modern lexical evidence are interpreted by the author as clear signs that the notion of a simple north-south divide should be replaced by that of fluid linguistic connections.

Christian Mair (269–283) challenges the traditional view of dialect from a chronological point of view. With the examples of Jamaican Creole and Nigerian Pidgin he demonstrates that in our time not only lexical, but also phonological and morphosyntactic features of originally marginal territorial varieties can spread into standard varieties whose speakers had no physical contact with those of the donor varieties. He identifies speaker mobility, pop-cultural movements, vernacular world literature and media exposure as the channels which help exterritorialize and globalize features of vernaculars. As a consequence of this new development he suggests a new concept of World English as "a pool of standard and non-standard features of varying and fuzzy regional reach" (277).

It was a very good decision of the editors to provide a list of abbreviations at the beginning of the volume. The idea was probably that it would guarantee that the same abbreviations were used for the historical periods of English and for the corpora, many of which were used by several authors, and that it would help identify the meaning of less well-known acronyms. It remains a riddle to me why the same abbreviations are introduced again in the individual papers. One might also have considered a similar condensation of information by bringing together the references of all papers in a single bibliography. This would have avoided reoccurring titles. Among the more often listed references I noticed Markus, Upton & Heuberger (2010) in four, and Beal (2004) and Mustanoja (1960) in three bibliographies each. There are also several printing errors, but they are unavoidable – as every author knows – and none of them really blurs

the meaning of the text. The book makes therefore profitable reading-matter for all anglicists who are interested in what is going on in diachronic English corpus linguistics.

—————

Lilo Moessner, RWTH Aachen
E-mail: moessner@anglistik.rwth-aachen.de