## Introduction: Corpora and the changing society

Paula Rautionaho[1], Arja Nurmi[2] and Juhani Klemola[2]

[1]University of Eastern Finland / [2]Tampere University

This volume presents 11 papers from the 39th ICAME[1] conference organized in Tampere in 2018, focussing on the theme of the conference, corpus linguistics and the changing society. It is a truism that both language and society are constantly changing. It is also commonly accepted that changes in society are reflected as changes in language and vice versa. When new technologies are introduced, we gain new vocabulary to discuss them. When our methods of communication change, so do our linguistic means of communicating; when the ideas and phenomena change and evolve, so does the language used to describe them. Words change their meaning as the entity they denote changes. Increasing gender equality, for instance, has led to changes in vocabulary referring to men and women (e.g. *chairman* vs. *chairperson*). Grammatical change may be more subtle, but nonetheless able to reflect cultural change; the decline in frequency of the English core modals and the concomitant increase of semimodals has been

---

[1] ICAME is the International Computer Archive of Modern and Medieval English, an international organization of linguists and information scientists working with English machine-readable texts. See http://clu.uni.no/icame/ for more information.

linked to the process of democratization. While we sometimes think of social and societal change as phenomena of the present day and language change as something only observed in the past, both are and have been present for as long as there have been languages and societies.

This volume approaches both patterns of change, those taking place on the level of language and those where social and linguistic change are reflected in each other. All papers included deal with English, choosing anything between a long diachrony from Old English to the present, or a timespan of a few decades in the immediate past. Two unifying threads can be seen again and again in the studies presented here: the ever-increasing availability of large diachronic corpora and the growing sophistication of corpus-linguistic methodology. In true ICAME tradition, the authors of the papers are aware of the potential pitfalls of the data and methods they use, and evaluate both as part of their linguistic inquiry into the phenomena they are studying.

The studies included in this volume approach the connections between language change and societal change with the tools offered by corpus linguistics. While the approach is by no means a new one, the advances in corpus methods and the availability of new datasets have made it possible to ask new questions and expand our understanding of different aspects of language and language change. The increasingly sophisticated statistical methods used to explore modern, extensive corpora allow for insights into questions previously studied with the help of small corpora, while at the

same time entirely new questions can be explored with the help of these tools. The cross-fertilization of cultural studies and corpus linguistics, we hope, is beneficial for both parties, corpus linguists and researchers of cultural change alike. Corpus linguistics offers a vast array of materials and methods to investigate cultural and societal change, while cultural studies provide the theoretical background on which to build our research. The studies included in the present volume showcase the potential avenues of combining changing language and changing societies.

The main body of the volume is divided into two parts, reflecting the two main approaches taken by the authors; some focus on the changing society, while others focus on the changing language. Part 1, *Changing Society*, consists of five papers which explicitly discuss the connections between changes in society and the ways in which this is reflected in language change and language use. The papers also raise questions regarding the reliability of corpus-based data and corpus methodology in identifying such connections.

**Martin Hilpert** starts by asking a question that is very relevant to the present volume: do "frequencies of linguistic items reflect their occurrence in the real world?" – is there a correlation between how we use language and the real world? Detecting new correlations between linguistic and social phenomena is ever more tempting with the abundance of resources today, but Hilpert draws together some of the pitfalls in the analysis of (diachronic) corpus data. He then goes on to suggest measures to be taken in order to

avoid, for instance, reporting spurious correlations, and reminds the reader of the importance of theoretical considerations in setting up a study.

**Gavin Brookes** and **David Wright** examine right-leaning press representations of non-native English-speaking migrants living in the UK. They identify where and how these representations have changed, or remained stable, over the period 2005 to 2017. The "speak English" debate emerges as a dynamic issue; during the period of time under investigation, the attitudes towards non-native English-speaking people shifted from an economic burden to a threat to the success of the local native groups. Brookes and Wright argue that the way that the right-leaning media represents this group of people ultimately affects the public opinion and government policy regarding migrants in the UK.

**Antoinette Renouf** combines a large corpus of newspaper data with a test of corpus-linguistic tools for tracking lexical innovation through collocation. She highlights the potential advantages as well as the disadvantages of the corpus tools AVIATOR, ACRONYM, and APRIL as means of identifying semantic change in the meaning of words, particularly in the culturally, socially, and politically driven changes in recent decades. She tests five different words (*birther*, *normalis/zation*, *cougar*, *snowflake*, and *ghosting*) which have gained new meanings and discusses the ways of identifying these meanings and the pitfalls of semi-automatic identification.

**Gerold Schneider** uses the topics of poverty and industrial revolution as a case study to show how societal and linguistic changes can be detected

using historical corpora. Building up a repository of statistical approaches, for instance topic modelling, Schneider shows how poverty, industrial revolution, and urbanization are associated with one another through, for instance, the associations of war, religion, and family.

Taking a collocational approach, **Maura Ratia** explores the changing attitudes of medical practitioners towards their patients in the years 1500 to 1800. By investigating the collocates of the words *patient* and *patients*, Ratia is able to detect a growing concern of the patients' well-being and suffering in the 18th century.

Part 2 of the volume, *Changing Language*, consists of six papers focusing on individual linguistic features and their development during a specific historical period, ranging from Old English to the 21st century. There is a particular focus on various types of adverbials, but changes and variation in the verb phrase are also discussed.

**Karin Aijmer** investigates the meanings and functions of the intensifier *absolutely* in recent British English. Her results show that the developments undergone by *absolutely* can be described as a trajectory from degree modifier to emphasizer and to discourse marker. A sociolinguistic account of the variables age and gender indicate that, perhaps contrary to expectations, it is the older speakers and male speakers who use *absolutely* more frequently than younger or female speakers.

Focusing on two death-related intensifiers, *deadly* and *mortal*, **Zeltia Blanco Suaréz** tracks their history from Old English to Present-day English.

A collocational analysis reveals that these intensifiers have evolved from their original literal meaning ('cause death') to the semantically bleached meaning ('very'). Blanco Suaréz concludes that *deadly* and *mortal* may well be victims of fashion: since Late Modern English they significantly decreased in frequency, and were replaced by other more powerful intensifiers.

**Yoko Iyeiri** focuses on adverbs ending in *-ingly*, tracking their history from Late Middle to Early Modern English, and finds that not only their frequency and types, but also functions have expanded. Iyeiri argues that the lexicalization of the *-ingly* adverbs may have been propelled by the creativity of individual writers in adopting and forming new adverbs. Finally, the so-called "Harry Potter adverbials" seem to have been in use earlier than what has been attested in existing research.

**Martin Schweinberger**'s study investigates the diachronic development of amplification of adjectives in American fiction data. His results indicate that *so* has replaced *very* as the dominant amplifier in predicative contexts in the 1990s. However, despite their general semantic similarity, *so* and *very* still exhibit distinct semantic profiles with respect to their semantic specialization. Schweinberger thus argues that it would be inaccurate to speak of full lexical replacement.

The topic of **Laurel Brinton**'s study is the development and functions of *that is not to say* (*that*), an introductory clause refuting an inference that could be drawn from the previous discourse. Combining diachronic and

synchronic approaches, Brinton tracks the frequency and pragmatic functions of this non-inference marker from the 16th century to the present day. The developments could be seen as the result of grammaticalization, but Brinton points out that some trends are actually more indicative of colloquialization.

Bringing the volume to its conclusion, **Turo Vartiainen** and **Mikko Höglund** focus on the historical development of a single verb, *sit*, taking stock of the variation between the simple and the SELF reflexive strategies in the Early Modern period, and the transitivization of *sit down* in Late Modern English. The study is one of the first to apply the Transitivity Hypothesis to historical data, and shows the usefulness of the hypothesis in explaining paths of transitivization in the evolution of English.

As the range of studies included in this volume illustrates, there are multiple ways in which to approach language and the changing society. The new insights gained through the application of up-to-date methodologies to modern, extensive datasets showcase the potential of this combination of viewpoints. Whether focussing on the interconnectedness of linguistic and social change or on the level of linguistics alone, the studies in this volume hopefully advance our understanding of not only the individual topics, but also more broadly the advantages of combining different outlooks and perspectives in new and innovative ways.