

# Lancaster Summer Schools in Corpus linguistics 2021:

## Open sessions

As part of the [Lancaster Summer Schools in Corpus Linguistics 2021](#), we are offering sessions, both lectures and a practical session, which are open to anybody who is interested in the topic. The sessions will be delivered through MS Teams an online platform used to deliver lectures and practical sessions. You can find the timetable for the sessions as well as links through which you can join the sessions below. **All times shown on this timetable are in UK time (BST).**

**[Dr Vaclav Brezina](#): *Numbers, graphs and party tricks: An introduction to statistics in corpus linguistics* (lecture)**

[LINK TO THE SESSION](#)

Numbers and graphs are two essential modes of scientific communication, complementing, evidencing and underscoring information communicated through text. Quantifying language through frequency analysis and visualizing of the results has a particular set of challenges, which we need to be aware of. This lecture offers an accessible introduction to statistics in corpus linguistics with a range of engaging examples of efficient statistical techniques and visualization of linguistic data. (Participants will also learn a few statistical party tricks in the process.)

**[Prof Elena Semino](#): *Corpus Linguistics and Health Communication* (lecture)**

[LINK TO THE SESSION](#)

Communication through language is central to the experience of illness and the provision of healthcare. This talk shows (a) why corpus linguistic methods are appropriate and needed in research on health communication, (b) how they can be employed, and (c) what they can contribute to healthcare research and practice. Examples will be drawn from communication about mental illness and about the Covid-19 pandemic.

**[Dr Anna Siyanova-Chanturia](#) (Victoria University of Wellington, New Zealand): *What corpus linguistics and psycholinguistics can tell us about the nature of multi-word expressions?* (lecture)**

[LINK TO THE SESSION](#)

Recent years have seen growing interest in the mechanisms behind the acquisition, processing and use of multi-word expressions (MWEs). MWEs encompass a large set of sequences above the word level, such as collocations (fast food), binomials (research and development), multi-word verbs (hold on), idioms (spill the beans), and so on. These sequences differ in many ways; however, what they all have in common is that they are highly familiar and, hence, highly predictable strings of language. In the present talk, I will provide an overview of the nature of MWEs from the perspectives of corpus linguistics and psycholinguistics.

**[Prof Jonathan Culpeper](#): Shakespeare's language and corpus approaches: Problems, solutions, insights (lecture)**

**[LINK TO THE SESSION](#)**

The kinds of difficulties that historical language presents corpus analysis are not generally different from those of present-day language. What differs is the degree of difficulty: historical language, especially from the more distant past, presents a chokingly thick concentration of every possible difficulty. I will touch on, amongst other things, spelling variation, words and lemmas and grammatical tagging. Along the way, I will provide insights into Shakespeare's language, including his supposedly massive contribution to new English vocabulary.

**[Dr Vaclav Brezina](#): #LancsBox: a software tool for corpus analysis of language data (practical session)**

**[LINK TO THE SESSION](#)**

This practical session will demo most recent features in the [#LancsBox package](#) including automatic production of research reports using the Wizard tool in #LancsBox. Participants will learn how to efficiently analyse and visualize linguistic data.

#LancsBox is a software package that incorporates a number of existing analytical techniques and adds new innovative methods that enable more efficient and sophisticated exploration of the data. #LancsBox can be used by linguists, language teachers, translators, historians, sociologists, educators and anyone interested in quantitative language analysis. It is free to use for non-commercial purposes and works with any major operating system. #LancsBox provides sophisticated analysis of large amounts of language data. It takes input files in any format (plain text, doc(x), pdf, odt, XML etc.) and processes data automatically adding part-of-speech annotation using the Treetagger (Schmitt 1995).

## Lancaster Symposium on Innovation in Corpus Linguistics 2021

*23 June 2021, 12.30pm – 4.30pm UK time*

Organised by the ESRC Centre For Corpus Approaches to Social Science ([CASS](#)), Lancaster University, UK. The symposium will take place online via MS Teams. This symposium is **free to attend**. Note: Lancaster Summer Schools in Corpus Linguistics 2021 participants do not need to register separately. They will be registered automatically.

Registration is required via the following link: <https://forms.office.com/r/kMpFAH6tu7>

**Programme [you can download the programme in pdf [here](#)]**

12.30 – 12.45 Introduction to the Symposium ([Vaclav Brezina](#))  
12.45 – 13.15 Corpus Linguistics and the Philosophy of Science ([Tony McEnery](#))  
13.15 – 13.45 Corpus Linguistics in Discourse Analysis ([Charlotte Taylor](#))  
13.45 – 14.15 Corpus Linguistics in Digital Humanities ([Michaela Mahlberg](#))  
14.15 – 14.30 Break  
14.30 – 15.00 Corpus Linguistics and Historical Sociolinguistics ([Terttu Nevalainen](#))  
15.00 – 15.30 Corpus Linguistics and Language Learning ([Fanny Meunier](#))  
15.30 – 16.00 Corpus Linguistics, Statistics and Tools Development ([Vaclav Brezina](#))  
16.00 – 16.30 General Discussion

The symposium offers a variety of perspectives on and applications of corpus linguistics. The focus of the symposium is on innovation in the field and the vision for the future. Each of the presenters was given the following four guiding questions to prepare their lecture.

1. Why is the corpus approach interesting for you personally?
2. Can you provide examples from your research that illustrate the value of using corpora?
3. Why does innovation matter in your research context?
4. What is your vision for corpora in your field in the future?

Participants will be able to post questions to the presenters during the symposium. Frequently asked questions will be addressed in the General discussion.

## Timetable of free events during Lancaster Summer Schools in Corpus Linguistics 2021

|                     | Monday<br>21 June                                                                                                                                | Tuesday<br>22 June                                                                                   | Wednesday<br>23 June                                                                                                                                                                                                  | Thursday<br>24 June                                                                                                                                                       |
|---------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 9.00<br>-<br>9.15   | <a href="#">Welcome</a><br><i>D. Gablasova<br/>et al.</i>                                                                                        |                                                                                                      |                                                                                                                                                                                                                       |                                                                                                                                                                           |
| 10.00<br>-<br>10.45 |                                                                                                                                                  |                                                                                                      |                                                                                                                                                                                                                       | <b>Lecture</b><br><a href="#">What corpus linguistics and psycholinguistics can tell us about the nature of multi-word expressions?</a><br><i>Anna Siyanova-Chanturia</i> |
| 12.30               |                                                                                                                                                  |                                                                                                      | <b>Lancaster Symposium on Innovation in Corpus Linguistics 2021:</b><br>Registration is required via the following link:<br><a href="https://forms.office.com/r/kMpFAH6tu7">https://forms.office.com/r/kMpFAH6tu7</a> |                                                                                                                                                                           |
| 15.00<br>–<br>15.30 |                                                                                                                                                  |                                                                                                      |                                                                                                                                                                                                                       | <b>Practical session</b><br><a href="#">#LancsBox: a software tool for corpus analysis of language data</a><br><i>Vaclav Brezina et al.</i>                               |
| 15.30<br>–<br>16.00 |                                                                                                                                                  |                                                                                                      |                                                                                                                                                                                                                       |                                                                                                                                                                           |
| 16.00<br>–<br>16.30 |                                                                                                                                                  |                                                                                                      |                                                                                                                                                                                                                       |                                                                                                                                                                           |
| 17:00<br>–<br>18:00 | <b>Lecture</b><br><a href="#">Numbers, graphs and party tricks: An introduction to statistics in corpus linguistics</a><br><i>Vaclav Brezina</i> | <b>Lecture</b><br><a href="#">Corpus Linguistics and Health Communication</a><br><i>Elena Semino</i> |                                                                                                                                                                                                                       | <b>Lecture</b><br><a href="#">Shakespeare's language and corpus approaches: Problems, solutions, insights</a><br><i>Jonathan Culpeper</i>                                 |