

Creating Corpus-Informed Materials for the English as a Foreign Language Classroom

Creating Corpus-Informed Materials for the English as a Foreign Language Classroom

A step-by-step guide for (trainee) teachers using online resources

ELEN LE FOLL (ED.)

*HANNA BAVENDIEK, SIMONE CAMPOS SILVA, JONAS
DAMMANN, GEESKE DÜNNHAUPT, ANNA FANKHAUSER,
MAGDALENA FUCHS, FABIAN GIESHOLT, MARIE GRÖNINGER,
LENA HAASE, LISA HÖRNSCHEMEYER, LENA JANZEN, JANA
KÜPERS, NELE LANGE, ELEN LE FOLL, ANNIKA LIEBELT,
TILZA MARIA MEISE-RECKEFUSS, HA MI NGUYEN, MARIE
NOTTMEIER, ALINA SOPHIE PETERS, NICOL SPERLING, RIEKE
STAHNKE, CARA SYMANZIK, KATINA VORBRINK, LARA
WARNECKE, AND MARLENA WEGESIN*



Creating Corpus-Informed Materials for the English as a Foreign Language Classroom by Elen Le Foll is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, except where otherwise noted.

Contents

How to use this book	1
<i>A quick guide to this Open Educational Resource</i>	
<i>Book structure</i>	1
<i>What is a corpus?</i>	2
<i>Get in touch!</i>	3
Corpus tools	4
<i>An overview of the tools and corpora used in this e-book</i>	
<i>English-corpora.org (formally BYU)</i>	4
<i>Sketch Engine</i>	5
<i>SKELL (Sketch Engine for language learning)</i>	6
About the project	7
<i>Co-creating an OER with student teachers to bridge the corpus linguistics research-practice gap</i>	
Elen Le Foll	
<i>1. Corpus linguistics and language teaching</i>	7
<i>2. The research-practice gap</i>	8
<i>3. The development of a project-based seminar</i>	9
<i>4. Challenges and lessons learnt</i>	11
<i>5. Learner autonomy</i>	13
<i>6. Conclusion and outlook</i>	14
<i>7. References</i>	15
Part I. Lesson ideas: Primary and lower secondary school	
1. "How's the weather?": Talking about the weather in the primary EFL classroom	18
Jana Küpers and Nele Lange	
<i>1 Introduction and rationale</i>	18
<i>2 Corpus, tools and methods</i>	19
<i>3 Step-by-step guide</i>	19
<i>4 Options and further ideas</i>	31
<i>5 Caveats and limitations</i>	32
<i>6 Conclusion</i>	33
<i>7 Resources and references</i>	33

2. "How do I get there?": Travelling by various means of transport	35
Magdalena Fuchs and Annika Liebelt	
1 <i>Introduction and rationale</i>	35
2 <i>Corpus, tools and methods</i>	36
3 <i>Step-by-step guide</i>	37
4 <i>Lesson plan</i>	45
5 <i>Options and further ideas</i>	50
6 <i>Conclusion</i>	54
7 <i>Resources and references</i>	54
3. "Let's have a party!": Teaching students how to express, accept and refuse invitations	56
Selke Tombrägel	
1 <i>Introduction and rationale</i>	56
2 <i>Corpus, tools and methods</i>	56
3 <i>Step-by-step guide</i>	57
4 <i>Options and further ideas</i>	69
5 <i>Caveats and limitations</i>	71
6 <i>Conclusion</i>	72
7 <i>Resources and references</i>	72
 Part II. Lesson ideas: Upper secondary school	
4. Differences in the grammatical structures of 'tell' and 'say'	75
<i>A data-driven learning approach</i>	
Geeske Dünnhaupt and Lena Haase	
1 <i>Introduction</i>	75
2 <i>Corpus/corpora, tools and methods</i>	76
3 <i>Step-by-step guide</i>	77
4 <i>Options and further ideas</i>	84
5 <i>Conclusion</i>	85
6 <i>Resources and references</i>	85
7 <i>Appendix</i>	85
5. Using corpora in the EFL classroom for creative writing	93
<i>The case of descriptive adjectives</i>	
Jonas Dammann	
1 <i>Introduction</i>	93
2 <i>Tools and methods</i>	94
3 <i>Step-by-step guide</i>	94
4 <i>Options and further ideas</i>	98
5 <i>Conclusion</i>	98
6 <i>Resources and references</i>	98

6. How to do data-driven learning (DDL) with secondary school pupils to improve their writing	99
Katharina Busmann	
1 <i>Introduction and rationale</i>	99
2 <i>Corpus, tools and methods</i>	100
3 <i>Step-by-step guide</i>	100
4 <i>Options and further ideas</i>	105
5 <i>Caveats and limitations</i>	112
6 <i>Conclusion</i>	113
7 <i>Resources and references</i>	113
8 <i>Appendix</i>	114
7. Teaching neologisms using corpus tools	116
Nicol Sperling and Sarah Winkelmann	
1 <i>Introduction and rationale</i>	116
2 <i>Outline and objectives</i>	116
3 <i>Corpus, corpus tools and methods</i>	117
4 <i>Teacher preparation</i>	117
5 <i>Lesson plan</i>	124
6 <i>Options and further ideas</i>	129
7 <i>Caveats & limitations</i>	130
8 <i>Conclusion</i>	131
9 <i>Resources and references</i>	132
10 <i>Appendix</i>	133
8. Commas in argumentative writing	138
Marie Nottmeier, Alina Sophie Peters, and Lara Warnecke	
1 <i>Introduction and rationale</i>	138
2 <i>Corpus, tool and methods</i>	139
3 <i>Step-by-step guide</i>	139
4 <i>Options and further ideas</i>	158
5 <i>Caveats and limitations</i>	159
6 <i>Conclusion</i>	159
7 <i>Resources and references</i>	160
8 <i>Appendix</i>	160

Part III. Lesson ideas: Content and Language Integrated Learning (CLIL) at secondary school

9. Teaching geography in English: How to describe a graph	165
Hanna Bavendiek and Fabian Giesholt	
1 <i>Introduction and rationale</i>	165
2 <i>Corpus, tools and methods</i>	166
3 <i>Step-by-step guide for Teachers</i>	167
4 <i>Options and further ideas</i>	172
5 <i>Caveats and Limitations</i>	172
6 <i>Conclusion</i>	173
7 <i>Resources and references</i>	173
10. "Black Lives Matter": Collecting vocabulary to improve intercultural communicative competence	174
Tilza Maria Meise-Reckefuß	
1 <i>Introduction and rationale</i>	174
2 <i>Corpus, tools and methods</i>	175
3 <i>Step-by-step guide: Teacher Preparation</i>	175
4 <i>Lesson plan</i>	190
5 <i>Worksheet</i>	191
6 <i>Options and further ideas</i>	199
7 <i>Conclusion</i>	200
8 <i>Resources and references</i>	200
11. Learning words and collocations independently: Rising temperatures and climate science	202
Lisa Hörnschemeyer and Lena Janzen	
1 <i>Introduction and rationale</i>	202
2 <i>Corpus, tools and methods</i>	203
3 <i>Step-by-step guide</i>	203
4 <i>Follow-up activities and further ideas</i>	213
5 <i>Conclusion</i>	213
6 <i>Resources and references</i>	213
7 <i>Appendix</i>	214

12. Act responsible, think sustainable – Expanding vocabulary to write about sustainability	217
Rieke Stahnke, Marlena Wegesin, and Lara Brede	
1 <i>Introduction and rationale</i>	217
2 <i>Corpus, tools and methods</i>	218
3 <i>Step-by-step guide</i>	218
4 <i>Options and further ideas</i>	237
5 <i>Caveats and limitations</i>	238
6 <i>Conclusion</i>	238
7 <i>Resources and references</i>	238
8 <i>Appendix</i>	239
13. Building a corpus to create materials for a debate on pollution	243
Marie Gröninger and Ha Mi Nguyen	
1 <i>Introduction and rationale</i>	243
2 <i>Corpus, tools and methods</i>	244
3 <i>Step-by-step guide</i>	244
4 <i>Options and further ideas</i>	254
5 <i>Caveats and limitations</i>	254
6 <i>Conclusion</i>	255
7 <i>Resources</i>	255
8 <i>References</i>	255
9 <i>Appendix</i>	256

Part IV. Lesson ideas: Vocational education and English for Specific Purposes (ESP)

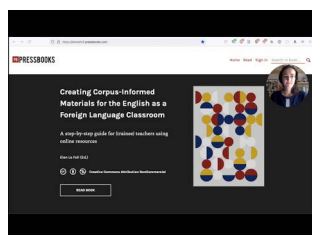
14. Developing ESP learning activities for vocational education with a custom virtual corpus	261
<i>Finding and exploiting topic-specific adjectives</i>	
Cara Symanzik	
1 <i>Introduction and rationale</i>	261
2 <i>Corpus, tools and methods</i>	262
3 <i>Step-by-step guide</i>	262
4 <i>Options and further ideas</i>	268
5 <i>Caveats and limitations</i>	268
6 <i>Conclusion</i>	269
7 <i>Resources and references</i>	269
8 <i>Appendices</i>	270

15. Exploring hospital interactions with Grey's Anatomy	272
<i>Building and exploiting a TV series corpus for nursing vocational schools</i>	
Simone Campos Silva	
1 <i>Introduction and rationale</i>	272
2 <i>Corpus, tools and methods</i>	273
3 <i>Step-by-step guide for teachers</i>	273
4 <i>Options and further ideas</i>	282
5 <i>Caveats and limitations</i>	283
6 <i>Conclusion</i>	283
7 <i>Resources and references</i>	283
8 <i>Appendices</i>	284
16. Creating teaching materials for nursing schools using medical TV series	288
Katina Vorbrink and Anna Fankhauser	
1. <i>Background</i>	288
2. <i>Methodology and procedure</i>	294
3. <i>Corpus-based teaching materials for the subject 'English / Communication'</i>	296
4. <i>Conclusion</i>	319
<i>References</i>	320
Appendix: Additional resources	324
Elen Le Foll	
1 <i>Online corpus tools and corpora</i>	324
2 <i>Free downloadable corpus tools for offline corpus work</i>	325
3 <i>Additional free online resources for teachers and learners</i>	325
4 <i>Recommended books</i>	326

How to use this book

A quick guide to this Open Educational Resource

This **Open Educational Resource (OER)** aims to empower English teachers from across the globe to design their own, authentic, corpus-based lessons by showcasing a range of ideas for creating **corpus-informed teaching materials** using online resources.



A YouTube element has been excluded from this version of the text. You can view it online here:
<https://elenlefall.pressbooks.com/?p=4>

Pre-service trainee teachers from **Osnabrück University** (Germany) contributed the chapters as part of three English Education Masters of Education seminars taught by **Elen Le Foll**. The introductory chapter “About the project” outlines the rationale and development of the project and discusses how various challenges were overcome.

The remaining **Lesson Ideas** chapters were contributed by students as part of their seminar coursework. These are organised according to the school type for which they were developed.

Book structure

- How to use this book
- Corpus tools
- About the project

The **Lesson Ideas** are grouped in four parts:

- **Part I** is dedicated to corpus-informed lesson ideas for **primary and lower secondary schools**.
- **Part II** showcases corpus-informed lesson ideas for **upper secondary schools**.
- **Part III** explores the use of corpora in **Content and Language Integrated Learning (CLIL)** and **bilingual secondary education**.

- **Part IV** presents corpus-informed lesson ideas for **English for Specific Purposes (ESP)** and **vocational education**.

A curated list of links to additional resources may be found in the Appendix.

For teachers and teacher trainees entirely new to corpora, we recommend selecting one or two chapters of interest and following the **step-by-step instructions** in order to recreate the corpus-informed materials proposed by the chapter authors. As you work your way through these, you will find that the various ideas and methods outlined in all the chapters can easily be transferred to an infinite range of different language foci, topics, and educational contexts.

What is a corpus?

A **corpus** (plural: **corpora**) is a principled, digital collection of texts. Texts can refer to **written** language (e.g., newspaper articles, blog posts, novels), **spoken** language (e.g., transcripts of conversations, debates, live TV shows) and **pseudo-spoken** language (e.g., film and TV series transcripts). It is important to note that the texts of a corpus were originally spoken or written for a **genuine, real-world communicative purpose**. In other words, unlike most texts found in textbooks, they were not created specifically for language teaching or learning purposes and can hence be considered to represent **authentic** sources of language.

In addition, most of the chapters include **worksheets** that can be downloaded as individual PDFs in just one click. Thus, this book also provides a low-threshold introduction to working with corpus-informed materials for teachers with no previous knowledge of corpora. It is hoped that the experience of using these “oven-ready” corpus-informed materials, which require little to no preparation time, will encourage teachers to subsequently invest time in working through a selection of the book’s chapters in order to, in due course, be able to pick their own ingredients and create entirely new and delicious corpus-informed dishes!

If you already have some experience using corpora for either research or teaching, you may immediately start using this book as a **bank of ideas and resources** to inspire your own recipes – tailoring your DIY materials to the specific needs and interests of your students.

Each chapter, or recipe, has a different focus which may be lexical, grammatical, or phraseological, and focuses on a different set of language and/or interdisciplinary skills. The chapters are all similarly structured. The chapter contributors begin by describing their lesson’s learning objectives and outlining the rationale for their choice of topic, corpus, and corpus tool. They then guide the reader through all the necessary steps to create their proposed corpus-informed materials with clear, tutorial-like and illustrated step-by-step instructions. In many instances, the authors also provide instructions for their lesson tasks, as well as (possible) solutions. At the end of each chapter, you will also find additional options and ideas to expand or adapt the proposed lesson to the taste buds of your students.

By publishing this e-book as a free, **Open Educational Resource** (OER), we hope to make the use of **corpus linguistics** in foreign language education accessible to as many teachers as possible across the world. All the chapters make use of corpora and corpus tools that are either **freely available** to all, or accessible with a free trial account or access to the European ELEXIS programme. Details of how to connect to and/or register to the various online corpus platforms are laid out in the respective chapters.

All the chapters are licensed under the Creative Commons Attribution-NonCommercial 4.0 International License, which means that they can be freely adapted, copied and distributed for non-commercial purposes as long as the original source and authors are cited. The full OER may be cited as follows:

Le Foll, Elen (Ed.). (2021). *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom. A step-by-step guide for (trainee) teachers using online resources* (Third Edition). Open Educational Resource. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0. DOI: 10.5281/zenodo.4992504.

Editable XML and **HTML versions** of this e-book can be downloaded from this project's Zenodo repository. A **PDF version** with hyperlinks is also available from the same repository for readers with unreliable internet connections – though we recommend using the **web-book version** available on [pressbooks.com](https://elenlefol.pressbooks.com).

Get in touch!

We would be delighted to hear about how you have used this book and whether it has inspired you to use corpora in your own English language teaching practice, so please **get in touch** via e-mail (elefol@uos.de) or on Twitter (@ElenLeFoll)! We also very much **welcome your suggestions** to improve this evolving and dynamic resource.

Elen Le Foll & student teachers from Osnabrück University

Corpus tools

An overview of the tools and corpora used in this e-book

The chapters of this book make use of different corpus tools and functions. If you are interested in learning how to use a specific tool or would like to use a specific corpus, you may find this list of the chapters sorted by the tools and corpora they rely on helpful.

English-corpora.org (formally BYU)

- Free basic access to all users and corpora (with a limited, but usually sufficient, number of queries).
- Registration required.
- Option to upgrade to premium individual or academic licenses.

With the Corpus of Contemporary American English (COCA):

- How to do data-driven learning (DDL) with secondary school pupils to improve their writing by **Katharina Busmann**
- Learning words and collocations independently: Rising temperatures and climate science by **Lena Janzen** and **Lisa Hörnschemeyer**
- Act responsible, think sustainable – Expanding vocabulary to write about sustainability by **Rieke Stahnke, Marlena Wegesin, and Lara Brede**

With the British National Corpus (BNC1994):

- Differences in the grammatical structures of 'tell' and 'say' by **Geeske Dünnhaupt** and **Lena Haase**

With the Corpus of American Soap Operas:

- "How's the weather?": Talking about the weather in the primary EFL classroom by **Jana Küpers** and **Nele Lange**

- “Let’s have a party!”: Teaching students how to express, accept and refuse invitations by **Selke Tombrägel**

With the News on the Web (NOW) corpus:

- Teaching neologisms using corpus tools by **Sarah Winkelmann** and **Nicol Sperling**
- Building a corpus to create materials for a debate on pollution by **Marie Gröninger** and **Ha Mi Nguyen**
- “Black Lives Matter”: Collecting vocabulary to improve intercultural communicative competence by **Tilza Maria Meise-Reckefuß**

With the Wikipedia Corpus:

- Developing ESP learning activities for vocational education with a custom virtual corpus by **Cara Symanzik**

Sketch Engine

- Limited free trial account (30 days).
- Registration required.
- Many EU higher education institutions have access via the EU-funded ELEXIS project.

With the family of EnTenTen corpora:

- Teaching geography in English: How to describe a graph by **Hanna Bavendiek** and **Fabian Giesholt**
- “How do I get there?”: Travelling by various means of transport by **Annika Liebelt** and **Magdalena Fuchs** (also uses SKELL)
- Using corpora in the EFL classroom for creative writing by **Jonas Dammann**

With the British National Corpus (BNC1994):

- Commas in argumentative writing by **Marie Nottmeier, Alina Sophie Peters** and **Lara Warnecke**

With instructions to create a custom corpus from web documents:

- Exploring hospital interactions with Grey's Anatomy by **Simone Campos Silva**

SKELL (Sketch Engine for language learning)

- Free and no registration required.
- Relies on a large corpus of web English.
- Also available for a number of other languages, e.g. German, French, Polish...

SKELL only gives access to one corpus:

- “How do I get there?": Travelling by various means of transport by **Annika Liebelt** and **Magdalena Fuchs**

Of course, there are many more tools and corpora also worth exploring...

About the project

Co-creating an OER with student teachers to bridge the corpus linguistics research-practice gap

ELEN LE FOLL

The practice of ELT (English Language Teaching) to date, at least, seems to be largely unaffected by the advances of corpus research, and comparatively few teachers and learners know about the availability of useful resources and get their hands on corpus computers or concordancers themselves. (Römer 2010: 18)

The development of this resource was spurred on by the observation that many corpus linguists have made: namely that the much-awaited “corpus revolution” has yet to have widely reached language teachers and learners outside of higher education contexts (e.g., McCarthy 2008; Chambers 2019). In this introductory chapter, I briefly outline the rationale and genesis of the project that led to the creation of this Open Educational Resource (OER). It was co-written with student teachers from Osnabrück University (Germany) as part of three iterations of a project-based seminar taught by the author between September 2019 and March 2021.

1. Corpus linguistics and language teaching

From the outset, applications of corpus linguistics to language teaching have been shown to be highly congruent with contemporary, evidence-based insights into second language acquisition (SLA). Rather than being made up of individual words strung together according to strict, grammatical rules, corpus linguistics has shown that language consists of vast networks of more or less fixed, frequently occurring patterns. In line with communicative approaches to language teaching, corpus linguistic methods help reveal the actual usage patterns and frequencies of real language as spoken or written in natural, real-world contexts.

Language learners can be explicitly taught these lexico-grammatical patterns or, as in Data-Driven Learning approaches (DDL; e.g., Boulton & Tyne 2013; Leńko-Szymańska & Boulton 2015), be encouraged to discover them for themselves. In a recent meta-analysis, Boulton & Cobb (2017) summarised the results of 64 experimental and quasi-experimental studies on the effectiveness of using corpus linguistics for second language (L2) learning or use, concluding that such approaches, on average, yield promisingly large positive effects ($d = 0.95$ for control/experimental group comparisons and $d = 1.50$ for pre/post-test designs).

Furthermore, allowing students to interact hands-on with corpora fosters interdisciplinary skills, such as digital skills and data literacy, and boosts learner autonomy. Not to be neglected either are the positive effects of successful teacher-corpus interactions. Even basic corpus literacy enables teachers to check their language intuition in an empirical manner, on the basis of authentic data. Thus, this project's foremost aim was to empower (trainee) teachers to create their own corpus-informed teaching materials, many of which also encourage data-driven learning and student-corpus interactions.

All four components of corpus literacy summarised by Callies (2019: 247, see also Mukherjee 2006: 14) are targeted:

1. Understanding basic concepts in corpus linguistics: What is a corpus and what types of corpora are available and how? What can you do – and cannot do – with a corpus?
2. Searching corpora and analysing corpus data by means of corpus software tools, e.g. concordancers: What is corpus software and how can it be used to search a corpus? How can corpus output be analysed?
3. Interpreting corpus data: How may general trends in language use/change be extrapolated from corpus data?
4. Using corpus output to generate teaching material and activities: How can you make use of corpus material for teaching purposes?

In the spirit of this OER, and of “open pedagogy” more generally (Willey & Hilton 2018), particular emphasis is placed on the use of freely accessible resources.

2. The research-practice gap

Whilst the value of corpora in foreign language teaching has long been understood, researched and documented (e.g., O’Keeffe, McCarthy & Carter 2007; Boulton & Cobb 2017; Pérez-Paredes 2019a), the “large gap between the wealth of applied corpus-linguistic research and the teaching practice” (Mukherjee 2004: 247; see also Chambers 2019; Pérez-Paredes 2019b) is still a reality. In school English as a Foreign Language (EFL) contexts, in particular, the “need to convince practising teachers to use corpora and concordances in the classroom” (Römer 2006: 129) has yet to have been met. To close this gap, a number of researchers have pointed to the centrality of teacher training (e.g., Mukherjee 2004; Hüttner, Smit & Mehlmauer-Larcher 2009; Breyer 2009). However, as Leńko-Szymańska (2014: 261) notes “[u]nfortunately, there are only a few books which serve as manuals for teachers in this area”. Things, however, do seem to be changing in this direction with some notable recent publications tackling the issue head-on: Eric Friginal’s “Corpus linguistics for English teachers: new tools, online resources, and classroom activities” (Routledge, 2019), Robert Poole’s “A guide to using corpora for English language learners” (Edinburgh University Press, 2018), and Dana Gablasova’s “Corpus for Schools” project (launched in 2017). A more comprehensive list of books and resources with full references can be found as an appendix to this e-book.

Aside from the resources from Gablasova’s Corpus for Schools project, which focus on spoken British English, the majority of these few existing resources are commercial book publications. This format bears two manifest disadvantages for teacher training purposes. First, if such books include practical information on how to use specific corpus tools, they very quickly become outdated (for instance, my students noticed that the host of changes made to english-corpora.org in 2020 made the step-by-step instructions of even very recent publications difficult, if not impossible, for corpus novices to follow). Secondly, but crucially, commercial book publications are only accessible to a small minority of privileged student teachers, teachers and teacher trainers.

With these aspects in mind, I decided to co-create, together with my students, an Open Education Resource to guide (trainee) teachers through the practical aspects of using corpora for language teaching using accessible, online resources.

3. The development of a project-based seminar

[T]here is, at present, a large gap between the wealth of applied corpus-linguistic research and the teaching practice in Germany which so far has only been affected to a very limited extent by this research. Closing this gap is a challenge to applied corpus linguists and, perhaps more importantly, to those who are involved in teacher training (both for trainee and qualified teachers). (Mukherjee 2004: 247)

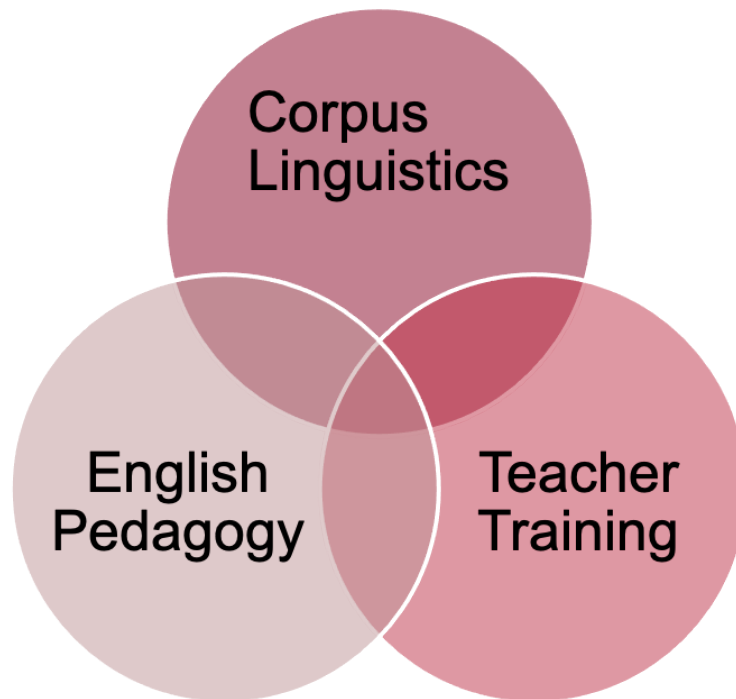


Fig. 1: The three core areas of the project-based seminar

Following Römer's (2010) call to focus more on language teachers' needs and inspired by previous such endeavours (Breyer 2009; Hüttner, Smit & Mehlmauer-Larcher 2009; Leńko-Szymańska 2014), I developed and subsequently evaluated a project-based seminar for M.Ed. students training to become English teachers (at primary, secondary and vocational school level) in Lower Saxony Germany (Le Foll 2020). The institutional constraints were: 13 weekly sessions of 90 minutes with ca. 30 students. The seminar was designed to convince pre-service student teachers of the value of corpus linguistics approaches to language teaching and learning and of its potential to boost learner autonomy. At the same time, the seminar aims to empower the aspiring teachers to be able to create corpus-informed materials autonomously using a range of tools and methods. The final project task consisted in co-writing and peer-reviewing a chapter for the present OER.

Seminar Learning Objectives

By the end of the semester, students should:

- Understand how Second Language Acquisition (SLA) research can inform materials design;
- Be able to evaluate existing English Language Teaching (ELT) materials;
- Be able to design corpus-informed ELT materials, incl. effective differentiated tasks, activities, instructions;
- Be confident users of basic corpus linguistic tools and methods;
- Be able to teach others how to design corpus-informed ELT materials;
- Be skilled at giving and learning from critical (peer) feedback.

The first few seminar sessions consisted of a hands-on introduction to the basic principles of corpus linguistics, including debunking some normative linguistic myths (cf. “surprise-the-teacher” modules suggested by Mukherjee 2004: 245) using online corpus tools (english-corpora.org, SketchEngine, BNClab, CQPweb, etc.) and a range of freely available corpora (COCA, BNC1994, Spoken BNC2014, Cambridge Learner Corpus, GloWbE, etc.). Following a Flipped Classroom (e.g., Reidsema et al. 2017) and Just-in-Time Teaching (e.g., Simkins & Maier 2010) approach, students “consumed” theoretical input (in the form of journal articles, book chapters and videos) in their own time and answered questions testing their understanding and asking them to reflect on what they had learnt. These answers informed the content of the synchronous class sessions.

German school EFL teachers are highly reliant on commercial textbooks and often reluctant to move away from them as they are perceived to embody the “one and only way to teach the curriculum” and certainly represent a “safe option”. To tackle this issue, the first phase in the seminar was to explore what makes good ELT materials and which SLA principles can be applied when evaluating textbook materials. Students overwhelmingly agreed that authentic (spoken) language usage and learner autonomy are two crucial aspects of language learning and that both are largely neglected in existing school EFL textbooks.

Based on this conclusion, I introduced the notion of corpus-based DDL in a hands-on approach: students tested their own (overwhelmingly non-native) language intuitions using corpora and, for greater ecological validity, completed DDL tasks designed to help them improve their own for academic writing (along others, with tasks adapted from Poole 2018). The advantages and limitations of data-driven learning (DDL) were discussed and students explored the alignment of the DDL approach with current SLA principles.

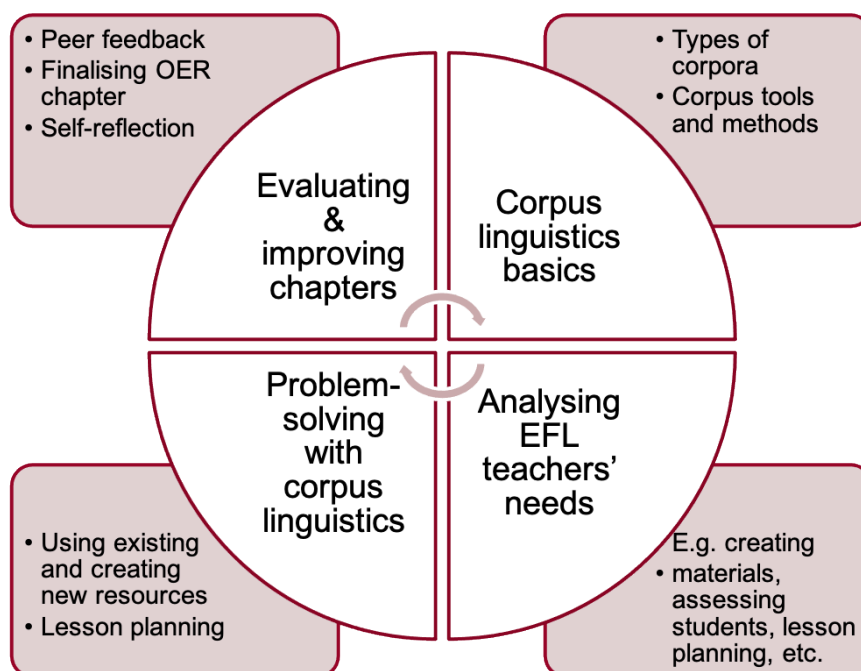


Fig. 2: Students' process for developing their co-authored OER chapters

Following a problem-solving approach, students were encouraged to think about the difficulties that ELF teachers typically face and, using a problem-solving approach, gathered ideas to develop appropriate corpus-based materials and tasks to help solve some of these issues. In lieu of an end-of-term examination or seminar paper, course assessment consisted in the contribution of a co-authored chapter to a “Practical Guide to Using Corpora for English as a Foreign Language Teachers”, to be published online (the attentive reader will note that the title of the OER has since changed!). The chapters submitted as coursework ranged from talking about breakfast (primary EFL), giving, accepting and refusing invitations in conversation (secondary EFL), teaching fractions (content and language integrated learning; CLIC), to talking to patients and next-of-kin about organ donation in hospital interactions (vocational education and training; VET).

Although the creation of the present OER was always at the heart of the seminar project, submitting co-authored coursework for publication was entirely optional and required the consent of all group members. In addition, I made clear that I would make a selection of the most suitable chapters. No additional credits could be obtained for taking part in the publication process. The contributing authors worked on improving their chapters on the basis of my and their peers' feedback, sometimes over several revision rounds, in their own time, often long after the seminar had ended. Given these conditions, it was very heartening to see that many of the authors of the selected chapters were very keen to contribute to the present publication.

4. Challenges and lessons learnt

Whilst very rewarding for all involved, designing and implementing such a project-based seminar is certainly not without its challenges. The present OER is the result of three iterations of the seminar, with three different groups of Master of Education students. Considerable adjustments to the seminar content and its implementation were made over the three semesters. In the following, I'd like to share some of aspects that initially did not go to plan, as well as the reasoning behind the adjustments made as the project unfolded and lessons were learnt.

The first iteration in the winter semester 2019/2020 was entitled “Corpus Linguistics and Language Teaching”. As the title suggests, the primary focus of the taught input was on corpus linguistics. Though student feedback was largely positive and the quality of the coursework overall very satisfactory, it was clear that some aspects required substantial changes. First, the primary focus of seminar needed to be shifted away from the theory and practice of using corpora, towards more theoretical and practical grounding in ELT materials design. Indeed, a number of students misinterpreted the seminar’s learning objectives as “learning all about the intricacies of as many corpora and corpus tools as possible”, thus largely relegating the pedagogical content of the seminar to the background. This resulted in student coursework that was more akin to (often well-designed) step-by-step guides to conducting corpus queries than about how to actually make pedagogically meaningful use of corpora in the classroom. As a result, I chose to entitle the second iteration of the seminar: “Designing and Evaluating Materials for Language Teaching”. Whilst student feedback from this second iteration was overall more positive than from the first, two students expressed the wish to have known before signing up the course that it involved designing ELT materials on the basis of corpus data. In order to avoid this misunderstanding and to make clear that corpus linguistics is a core part of the seminar, the third iteration was entitled: “Designing and evaluating corpus-informed language teaching materials”. Interestingly, a number of students later reported choosing this third seminar specifically because they wanted to learn more about how to use corpora.

A number of students from the first iteration of the “Corpus Linguistics and Language Teaching” seminar also reported that the wealth of available corpus tools and functions was overwhelming. Consequently, the second major change, which came hand in hand with the first, was to drastically reduce the course’s teacher-led input on corpus software, platforms and query functions. The first seminar covered english-corpora.org, SketchEngine, LancsBox, AntConc and CQPweb. The latter was only used to access the Spoken BNC2014, which was subsequently also uploaded onto SketchEngine so that future iterations did not need to include it. LancsBox was introduced together with a collaborative session in which the class created a higher education EFL/ESL learner corpus to which students contributed their own anonymised coursework from previous semesters. In practice, however, few students were willing to contribute any coursework to this toy learner corpus, which made it difficult to draw any meaningful conclusion from its analysis. A second activity required students to use AntCorGen to create their own discipline-specific academic corpus. These two activities proved difficult for many students. First, there were a number of technical difficulties. Although both software are freely accessible and work well cross-platform, a handful of students had seemingly never installed any piece of software before and/or only had access to computers for which they did not have installation rights. All in all, the number of hours I invested in one-to-one technical support proved to be unsustainable. In order to make the best use of the limited time available and ensure equal participation for all, I therefore decided that future iterations of the course would only include browser-based online tools. Another strong motivator for this decision was my wish for the seminar objectives to be as long-term and sustainable as possible. As many practicing teachers will know, installing software on school computers/tablets is, at best, a frustratingly laborious task and, often, quite simply mission impossible! Using (freely accessible) browser-based tools removes this first, practical barrier to using corpora in school contexts. Moreover, a number of students later reported querying online corpus tools on their mobile phones to find quick solutions to their own language questions, which, again, highlights the long-term sustainability of such an approach. In any case, I strongly believe that having acquired a sound knowledge of corpus skills in any of the commonly used online or offline corpus tools, motivated teachers will have no trouble finding their way around other corpus tools. Indeed, this was shown to be case with some of my students successfully choosing to use CQPweb, SKELL and the TED Corpus Search Engine for their final projects, in spite of these not having been introduced in any of the compulsory or optional course tasks. Equally, I always referred interested students to AntConc, LancsBox, and other free, offline software, whenever these were likely to be better suited for their project ideas.

In addition, the first iteration of the seminar included a few technology-adverse students who reportedly did not enjoy the hands-on computer-based work and did not respond to DDL as positively as the rest of the participants. In future iterations, I made sure to stress that corpus-informed materials can just as

well be paper-based and need not necessarily rely on hands-on learner-corpus interactions (cf. Boulton 2010 on paper-based DDL activities). It is difficult to judge whether any future technology-adverse students were warned by their peers not to join this course, or whether the subsequent versions of the seminar better addressed their concerns because, as a result of the COVID-19 pandemic, the second and third iterations of the seminar were taught online so that all students were, at any rate, forced to study online and work with their own devices. The online iterations of the seminar involved a combination of an elaborate asynchronous learning portfolio with individual and collaborative tasks and bi-weekly synchronous webinar sessions. Whilst this new seminar format required a lot of adaptations on my part, a number of students highlighted the advantages of the online course format for this project-based seminar:

Gerade hinsichtlich des behandelten Themas war es sehr praktisch, dass jeder an seinem eigenen Computer saß. Die Möglichkeit, Arbeitsprozesse mit einem Corpus erst über einen Screenshot vermittelt zu bekommen und es quasi simultan ausprobieren zu können, empfand ich als vorteilhaft (gegenüber herkömmlichen Seminarstrukturen) [Given the nature of the topics covered in the seminar, it was actually very handy that everyone was sitting in front of their own computers. Compared to the usual seminar structures, I found it advantageous to be taught how to use a corpus via screencast and be able to try it out for myself almost instantaneously.]” (anonymous feedback from student evaluation, summer 2020)

In spite of all the challenges, the majority of students reported a high degree of satisfaction with the course and especially with the outcome of the final project task:

Over the course of the last weeks, I started to really appreciate the power of corpora and I am now excited to do a corpus-based teaching unit with a class. (student self-reflection, winter 2019/2020)

Since I would like to become a teacher for secondary education, I can imagine using exactly the chapter my fellow students and I have written in school once. (student self-reflection, winter 2020/2021)

Generally, I missed the connection to my future teaching job in a lot of my other courses, so I really liked that this course’s final task was designing something that could actually be used in school one day. (student self-reflection, winter 2020/2021)

5. Learner autonomy

The seminars aimed to persuade student teachers of the value of corpus linguistics approaches to language teaching and their potential to boost learner autonomy. It empowered teachers to create corpus-informed materials autonomously and, via the creation and publication of the present OER, enabled them to become future corpus-informed materials design and data-driven learning (DDL) multipliers.

In the seminars, learner autonomy was tackled at two levels. First, all the participants were future English teachers and, although they are M.Ed. students, the majority of their university courses consist of frontal teaching and feature relatively few elements of self-regulated learning. Thus, this seminar also aimed to develop future teachers’ own learning autonomy by allowing them to explore, apply, reflect and exchange on various ELT and SLA principles. They learnt about creating and evaluating ELT materials that support learner autonomy through DDL. DDL was new to all participating student teachers and was, initially, met with some resistance. Given their own relative lack of experience with such approaches, this was to be expected. In the pre-intervention questionnaire, two-thirds of the students agreed or strongly agreed with the statement: “I have no idea what corpus linguistics is about”. Only a handful of participants reported already using corpora. All students responded negatively to the statement: “I know how to use corpora to

prepare classroom materials”. In light of these results, the following chapters, which were developed by the contributing student authors in just one semester, are even more impressive!

6. Conclusion and outlook

I am delighted that a small but highly motivated team of thirty students from the three iterations of this project-seminar has remained on board to see the project through to the publication stage. Although most students co-wrote their chapters in groups of two or three, five chapters were developed by students who had decided to work individually. I would like to thank all contributing authors, as well as all the students who participated in the seminars and who therefore also directly (e.g., through their peer review of the present chapters) and indirectly (e.g., through their course feedback which helped me improve future seminars) contributed to the successful completion of this project. Heartfelt thanks are also due to Tatjana Winter who, in her capacity as a student research assistant at the Institute of English and American Studies (IfAA) at Osnabrück University, formatted the chapters for publication on pressbooks.com, and to my PhD supervisor Dirk Siepmann without whose support this (very time consuming!) side project would not have been possible.

All that remains for me to write is that I hope that this OER will prove to be a useful, practical resource for EFL/ESL teachers from across the world to learn to design their own corpus-informed materials and draw inspiration from. It is also highly suitable for use as a textbook or complementary resource in both pre- and in-service teacher training programmes on corpus linguistics in language education.

All the chapters are licensed under the Creative Commons Attribution-NonCommercial 4.0 International License, which means that they can be freely adapted, copied and distributed for non-commercial purposes as long as the original source and authors are cited. **Editable XML** and **HTML versions** of this e-book can be downloaded from this project’s Zenodo repository. A **PDF version** with hyperlinks is also available from the same repository for readers with unreliable internet connections – though we recommend using the **web-book version** available on pressbooks.com.

The full OER may be cited as follows:

Le Foll, Elen (Ed.). (2021). *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom. A step-by-step guide for (trainee) teachers using online resources* (Third Edition). Open Educational Resource. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

Individual chapters should be referenced with the authors’ names and the full link to the chapter, e.g.,:

Nottmeier, Marie, Alina Sophie Peters & Lara Warnecke (2021). “Commas in argumentative writing”. In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. https://elenlefol.pressbooks.com/chapter/nottmeier_peters_warnecke. CC-BY-NC 4.0.

We welcome your feedback!

If you spot any errors, dead links, or have any other kind of feedback, do please get in touch via e-mail (elefoll@uos.de) or on Twitter (@ElenLeFoll). The contributing authors and myself very much look forward to hearing how you are using this e-book. Do drop us a note to tell us how you are using and/or adapting the resource for your teaching and learning context! We also very much **welcome your suggestions** to improve this evolving and dynamic resource.

7. References

- Boulton, Alex. 2010. Data-Driven Learning: Taking the Computer Out of the Equation: Data-Driven Learning. *Language Learning* 60(3). 534–572. <https://doi.org/10.1111/j.1467-9922.2010.00566.x>.
- Boulton, Alex & Tom Cobb. 2017. Corpus use in language learning: A meta-analysis. *Language Learning* 67(2). 348–393. <https://doi.org/10.1111/lang.12224>.
- Boulton, Alex & Henry Tyne. 2013. Corpus linguistics and data-driven learning: a critical overview. *Bulletin suisse de Linguistique appliquée* (Neuchâtel: Institut de Linguistique de l'Université) 97. 97–118. <https://hal.archives-ouvertes.fr/hal-01208263>.
- Breyer, Yvonne. 2009. Learning and teaching with corpora: reflections by student teachers. *Computer Assisted Language Learning* 22(2). 153–172. <https://doi.org/10.1080/09588220902778328>.
- Callies, Marcus. 2019. Integrating corpus literacy into language teacher education: The case of learner corpora. In S. Götz & J. Mukherjee (Eds.), *Studies in Corpus Linguistics* (Vol. 92, pp. 245–263). Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.201.12cal>
- Chambers, Angela. 2019. Towards the corpus revolution? Bridging the research–practice gap. *Language Teaching* 52(4). 460–475. <https://doi.org/10.1017/S0261444819000089>.
- Deutscher Anglistenverband & Deutsche Gesellschaft für Amerikastudien. 2009. Inhaltliche Anforderungen für Fachwissenschaft und Fachdidaktik in der Lehrerinnen- und Lehrerbildung, Studienfach Englisch (Positionspapier). Retrieved from http://www.anglistenverband.de/wp-content/uploads/LA-Englisch_Positionspapier.pdf
- Hüttner, Julia, Ute Smit & Barbara Mehlmauer-Larcher. 2009. ESP teacher education at the interface of theory and practice: Introducing a model of mediated corpus-based genre analysis. *System* 37(1). 99–109. <https://doi.org/10.1016/j.system.2008.06.003>.
- Le Foll, Elen. 2020. Development and Evaluation of a Corpus Linguistics Seminar in Pre-Service Teacher Training. Presented at the Teaching and Language Corpora Conference (TaLC) 2020 in Perpignan. <https://doi.org/10.13140/RG.2.2.10720.10249>. Video recording: <https://www.youtube.com/watch?v=PtgW5y-xFW8>.
- Leńko-Szymańska, Agnieszka. 2014. Is this enough? A qualitative evaluation of the effectiveness of a teacher-training course on the use of corpora in language education. *ReCALL* 26(2). 260–278. <https://doi.org/10.1017/S095834401400010X>.

- Leńko-Szymańska, Agnieszka & Alex Boulton (eds.). 2015. *Multiple affordances of language corpora for data-driven learning* (Studies in Corpus Linguistics (SCL) Volume 69). Amsterdam; Philadelphia: John Benjamins.
- McCarthy, Michael. 2008. Accessing and interpreting corpus information in the teacher education context. *Language Teaching* 41(4), 563–574. <https://doi.org/10.1017/S0261444808005247>.
- Mukherjee, Joybrato. 2004. Bridging the Gap between Applied Corpus Linguistics and the Reality of English Language Teaching in Germany. In U. Connor & T. Upton (eds.), *Applied Corpus Linguistics: A Multidimensional Perspective*, 239–250. Amsterdam: Rodopi.
- Mukherjee, Joybrato. 2006. Corpus linguistics and language pedagogy: The state of the art—and beyond. In Sabine Braun, Kurt Kohn & Joybrato Mukherjee (eds.), *Corpus technology and language pedagogy: New resources, new tools, new methods*, 5–24. Frankfurt: Peter Lang.
- O’Keeffe, Anne, Michael McCarthy & Ronald Carter. 2007. *From corpus to classroom: Language use and language teaching*. Cambridge; New York: Cambridge University Press.
- Pérez-Paredes, Pascual. 2019a. A systematic review of the uses and spread of corpora and data-driven learning in CALL research during 2011–2015. *Computer Assisted Language Learning* 1–26. <https://doi.org/10.1080/09588221.2019.1667832>.
- Pérez-Paredes, Pascual. 2019b. The pedagogic advantage of teenage corpora for secondary school learners. In Peter Crosthwaite (ed.), *Data-Driven Learning for the Next Generation*, 67–87. London; New York, NY: Routledge. <https://doi.org/10.4324/9780429425899-5>.
- Poole, Robert. 2018. *A guide to using corpora for English language learners*. Edinburgh: Edinburgh University Press.
- Reidsema, Carl, Lydia Kavanagh, Roger Hadgraft & Neville Smith. 2017. *The Flipped Classroom: Practice and Practices in Higher Education*. <https://doi.org/10.1007/978-981-10-3413-8>.
- Römer, Ute. 2006. Pedagogical applications of corpora: Some reflections on the current scope and a wish list for future developments. *Zeitschrift für Anglistik und Amerikanistik* 54(2), 121–134.
- Römer, Ute. 2010. Using general and specialized corpora in English language teaching: Past, present and future. In Mari Carmen Campoy Cubillo, Begoña Bellés Fortuño & María Lluïsa Gea-Valor (eds.), *Corpus-based approaches to English language teaching*, 18–38. London: Continuum.
- Simkins, S. & M. Maier. 2010. *Just-in-time teaching: Across the disciplines, across the academy* (New Pedagogies and Practices for Teaching in Higher Education Series). Stylus.
- Wiley, D., & Hilton, J. L. (2018). Defining OER-Enabled Pedagogy. *The International Review of Research in Open and Distributed Learning*, 19(4). <https://doi.org/10.19173/irrodl.v19i4.3601>

PART II

LESSON IDEAS: PRIMARY AND LOWER SECONDARY SCHOOL

Corpus-informed EFL lessons for primary and lower secondary school

- “How’s the weather?”: Talking about the weather in the primary EFL classroom by **Jana Küpers** and **Nele Lange**
- “Let’s have a party!”: Teaching students how to express, accept and refuse invitations by **Selke Tombrägel**
- “How do I get there?”: Travelling by various means of transport by **Annika Liebelt** and **Magdalena Fuchs**

1. "How's the weather?": Talking about the weather in the primary EFL classroom

JANA KÜPERS AND NELE LANGE

1 Introduction and rationale

The weather plays a central role in everyone's daily life and is therefore also an essential part of everyday English: it is a common topic for small talk, when talking about our holidays, and is also a crucial factor in our choice of clothing. Thus, weather is a topic everyone can refer to and that everyone can say something about. This is also why it is one of the first topics to be taught in English as a Foreign Language (ELF) classes at primary school. In this way, the topic "Weather" contributes to achieving the main objective of all EFL teaching: "to equip learners with competences that allow them to communicate with other speakers of the English language in a respectful manner" (Surkamp & Viebrock 2018: 21) and is thus also in line with the primary school curriculum of Lower Saxony, Germany (Niedersächsisches Kultusministerium 2016: 7).

However, the majority of EFL tasks about the weather in primary school textbooks often only require pupils to repeat, assign or classify pictures and/or single words, which does not help them develop communicative and intercultural competences. According to Thaler (2012: 84), one of the main disadvantages of tasks in EFL textbooks is that they can be repetitive and do not really activate pupils. In line with the core objectives of teaching English as a foreign language, the EFL materials and tasks created in this chapter therefore aim for a more effective and interactive way of learning English at primary school. This chapter provides teachers with a step-by-step guide on how to effectively make use of an appropriate corpus to compile idiomatic spoken language materials that enable primary school pupils to engage in (brief) authentic conversations about the weather in- and outside the classroom.

Outline and objectives

Target learners: Primary school pupils at the end of year 3 or (the beginning of) year 4

Teacher preparation time: 90 to 120 min (for three tasks)

Lesson time: 3x 45 min.

Learning objectives:

By the end of the lesson, pupils should be able to...

- answer questions about today's weather,
- ask others about today's weather,

- understand and use a range of adjectives typically used to describe the current weather situation in natural conversations,
- understand a simplified weather report.

2 Corpus, tools and methods

As even third grade primary school pupils are still in the process of acquiring and learning grammatical and orthographical features of the literacy system of their native language, the focus when teaching English as a foreign language in primary school should clearly be on oral skills (speaking and listening) (cf. Suhrkamp & Viebrock 2018: 26). The unit described in this chapter therefore mainly aims to improve pupils' speaking and listening skills. Accordingly, it is advisable to create teaching materials based on language data derived from a corpus of spoken English. The Corpus of American Soap Operas is particularly suitable for this purpose. Several scholars have highlighted the potential of soap operas for the acquisition of a foreign language as they introduce learners to very natural and authentic spoken-like language (Jones & Horak 2014: 12-13; Quaglio 2009: 148-149). The Corpus of American Soap Operas contains 100 million words from more than 22,000 transcripts of ten American soap operas aired between 2001 and 2012. Moreover, this corpus does not include any swear words and is therefore particularly suitable as a basis for the generation of teaching materials for younger learners. The corpus can be accessed freely on www.english-corpora.org. Although it is necessary to register to use this online platform, the creation of a free account (which limits the number of searches per day) is more than sufficient for our purposes and does not restrict the website's functions in any other way.

The data used for designing teaching materials can be derived with the help of the website's "Collocates"-function. Collocations are words that "occur most frequently and with statistical significance (i.e. not just by random occurrence) in the word's environment" (O'Keeffe, McCarthy & Carter 2007: 14). Both actively teaching collocations and making pupils aware of them is key to successful EFL learning, as learning collocations "contributes to the learner's ability to create associations between words and to place them meaningfully within various networks in relation to other words" (O'Keeffe, McCarthy & Carter 2007: 54). Therefore, it is argued that "learning the collocations of that language is not a luxury if anything above a survival level mastery of the language is desired, since collocation permeates even the most basic, frequent words" (O'Keeffe, McCarthy & Carter 2007: 60). Consequently, it can be seen as a crucial factor for the acquisition of authentic language skills that primary school pupils are given the opportunity to work with EFL materials which include collocations and that they are made aware of these collocations by both the materials and the teacher.

3 Step-by-step guide

3.1 Teacher preparation

3.1.1 Data collection – Conducting corpus queries

Step 1: Registration on www.english-corpora.org

First, go to www.english-corpora.org. Create a free account. Click on “My account” and then on “Register”.



Fig. 1: Registration on www.english-corpora.org

You will then have to fill in a registration form. After submitting this form, you will receive an e-mail which includes a confirmation link.

Please fill out the brief form below. Within one or two minutes, you will receive an email. Simply click on the link in that email, and you will be able to continue using the corpora. (Note: If you have already registered, please [log in](#) to see your profile)


Name	<input type="text"/> <input type="text"/> (e.g. Mary Smith)
Email address	<input type="text"/>
Password	<input type="password"/>
Country	-- SELECT--
Category	<input type="radio"/> University professor: languages / linguistics <input type="radio"/> Graduate student: languages or linguistics <input type="radio"/> University professor: not languages / linguistics <input type="radio"/> Teacher: not university; not graduate student <input type="radio"/> Graduate student: not languages or linguistics <input type="radio"/> Student (undergraduate) <input checked="" type="radio"/> Other
	3 RESEARCHER
	2 SEMI-RESEARCHER
	1 NOT RESEARCHER
	<input checked="" type="checkbox"/> I agree to the Terms and Conditions for this website
	<input type="button" value="SUBMIT"/> <input type="button" value="RESET"/> PROBLEMS ??

Fig. 2: Registration form on www.english-corpora.org

Step 2: Selecting a corpus

Next, on the home page, select a corpus from which you would like to draw data. For the unit described in this chapter, the Corpus of American Soap Operas is particularly suitable (see section 2).

The most widely used online corpora. [Overview](#), [search types](#), [looking at variation](#), [corpus-based resources](#).

The links below are for the online interface. But you can also  download the corpora for use on your own computer.











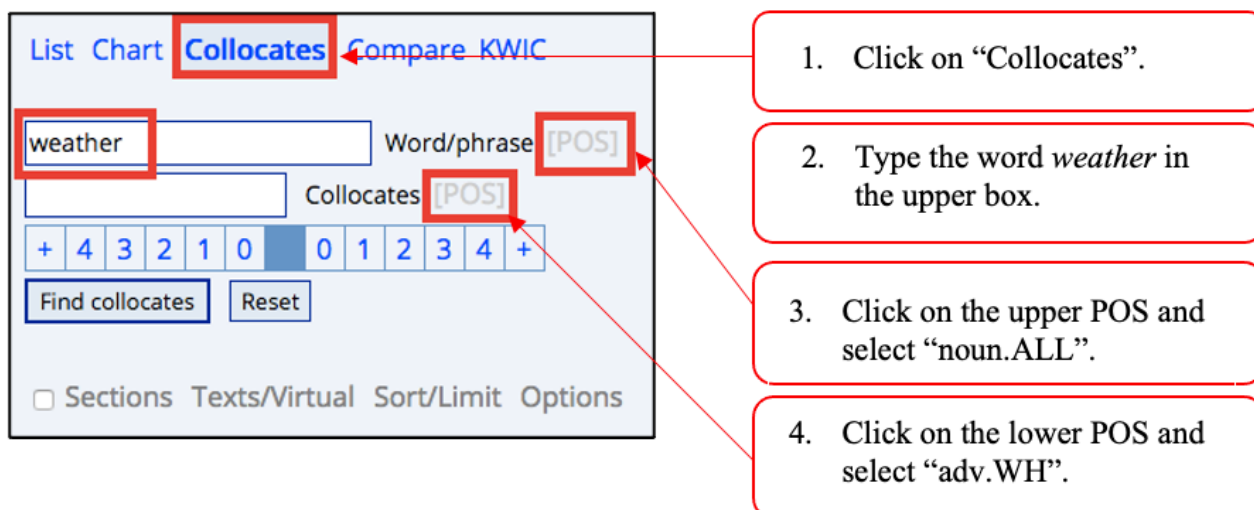
Corpus (online access)	Download	# words	Dialect	Time period	Genre(s)
iWeb: The Intelligent Web-based Corpus		14 billion	6 countries	2017	Web
News on the Web (NOW)		10.5 billion+	20 countries	2010- yesterday	Web: News
NEW The Coronavirus Corpus		468 million+	20 countries	Jan 2020- yesterday	Web: News
Global Web-Based English (GloWbE)		1.9 billion	20 countries	2012-13	Web (incl blogs)
Wikipedia Corpus		1.9 billion	(Various)	2014	Wikipedia
Corpus of Contemporary American English (COCA)		1.0 billion	American	1990- 2019	Balanced
Corpus of Historical American English (COHA)		400 million	American	1810-2009	Balanced
The TV Corpus		325 million	6 countries	1950-2018	TV shows
The Movie Corpus		200 million	6 countries	1930-2018	Movies
Corpus of American Soap Operas		100 million	American	2001-2012	TV shows

Fig. 3: Selecting a corpus

Step 3: Searching for question words – How to ask somebody about the weather

Now you can start searching for language data that can be used for designing your teaching materials. As the topic of the unit is “Weather”, it is advisable to firstly identify how to ask somebody about the weather. In order to do so, you need to use the “Collocates”-function. With the help of this function, you can determine question word(s) which occur together with the word *weather*. For this, you need to type *weather* in the upper box and click on “POS” (= part of speech), which appears in light grey letters next to this box. With regard to the lower box, you only need to click on “POS”, which again appears in light grey letters next to the lower box, and select “adv.WH” (Fig. 4). This way, you will obtain a list including all the adverbs that occur together with the noun *weather* and simultaneously function as question words.



The screenshot shows the 'Collocates' search interface. It features a top navigation bar with 'List', 'Chart', 'Collocates', and 'Compare KWIC'. Below this, there are two input boxes: 'Word/phrase' containing 'weather' and 'Collocates'. To the right of each box is a '[POS]' dropdown menu. Below the input boxes is a frequency scale from -4 to +4, with '0' in the center. There are 'Find collocates' and 'Reset' buttons. At the bottom, there are checkboxes for 'Sections', 'Texts/Virtual', 'Sort/Limit', and 'Options'. Four numbered instructions are overlaid on the interface:

1. Click on “Collocates”.
2. Type the word *weather* in the upper box.
3. Click on the upper POS and select “noun.ALL”.
4. Click on the lower POS and select “adv.WH”.

Fig. 4: How to fill in your search box

Last, you need to select the position in which the question word occurs within the sentence. As question words usually occur at the beginning of a question and thus in front of the word *weather*, you need to

select a number on the left side. Since one can assume that there are about two to three words between the question word and the word *weather*, it is advisable to click on number four. In the end, your search box should look like the one shown in Fig. 5:

Fig. 5: What your search box should look like at the end

If you click on “Find collocates”, a list with question words should appear on your screen (Fig. 6). The most frequent question word used before the word *weather* is *how*.

HELP	<input type="checkbox"/>	CONTEXT	FREQ
1	<input type="checkbox"/>	HOW	34
2	<input type="checkbox"/>	WHEN	7
3	<input type="checkbox"/>	WHY	1
4	<input type="checkbox"/>	WHERE	1
TOTAL			43

Fig. 6: List of question words that function as an adverb and occur in combination with the noun *weather*

If you click on the word *how*, a list appears, showing concrete example sentences in which this particular question word occurs in combination with the noun *weather* (Fig. 7):

CLICK FOR MORE CONTEXT					<input type="checkbox"/> [?]	SAVE LIST	CHOOSE LIST	CREATE NEW LIST	<input type="text"/>	[?]	SHOW DUPLICATES
1	2011	AMC	A	B	C	I'm not gon na let that happen. Colby: So, how's the weather in cuckooland these days, Annie? Annie: You may think I'm crazy --						
2	2011	GH	A	B	C	(Cell-phone-rings) Theo: Yes? Yes. See to it. So, how's the weather in Cajun country this morning? Did my grandson's caretaker say? Suzanne: You						
3	2010	AMC	A	B	C	intrude. Sit down. Heh heh heh! Madison: So, how was the weather ? Ryan: The weather? The weather. Actually, I don't even know						
4	2010	AMC	A	B	C	The weather? The weather. Actually, I don't even know how the weather was. I went from the airport to the hotel to the boardroom. I						
5	2009	GL	A	B	C	and hot. Of course, you know most coffee is. How's the weather out there? Olivia: Stop it. Cut the cute talk, okay. We						
6	2009	YR	A	B	C	listen to him, Kevin. He is not real. Chipmunk-voice How's the weather in la-la land, Pal? Amber: Ignore him. What is real is right						
7	2008	GH	A	B	C	I mean, there are plenty of other things to discuss. How about this weather we're having? No, seriously, there's politics. I mean,						
8	2008	GL	A	B	C	Yeah. Frank: Is she showing the house? Rick: No. How's the weather ? Coop: Yes, you want to talk about Springfield? The Mayor? Rick:						
9	2008	YR	A	B	C	Nikki: They tried, but all they got out of me was how nice the weather and food are in Brazil. Michael: Good. Okay, we're gon na						
10	2007	GL	A	B	C	well. I'm strong like you. Hello, sunshine. How's the weather in there? I bet you wonder why I'm here. I went to						
11	2007	GL	A	B	C	weeks. I mean, have you seen any good movies? How was the weather ? Any news about anything? Alan-Michael: Well, Josh Lewis is in jail.						
12	2006	AMC	A	B	C	a letter. A good-bye letter, to Tad. # Tad: How is the weather in the South of France, anyway? Krystal: Oh, come on. Are						
13	2005	AMC	A	B	C	just fine. Zach: Yeah, evidently, so will you. How's the weather up there? Kendall: Oh, well, it's not too cool. A						
14	2005	GH	A	B	C	Great. Cancel my 2:00. Woman: You're welcome. Louise: How's the weather ? Woman: Cloudy with a 50% chance of " don't bother me. "						
15	2005	GH	A	B	C	? Elizabeth: Yeah, you could say that. Jax: Yeah. How about this weather we've been having? You know what? Courtney: Oh, would you stop						
16	2005	GL	A	B	C	n't really mean that I want you? Like when we ask how's the weather it won't really mean that I am dying to be with you? Like						
17	2005	GL	A	B	C	Marina: France. # Buzz: Okay, if you say so. How is the weather in Paris? # Marina: Actually, we're still in the air, but						
18	2004	DAYS	A	B	C	be around people. I need to hear trivial conversation. " How's the weather ? " " How's your new house coming along? " Julie: I know						
19	2004	OLTL	A	B	C	's me. I just wanted to -- I just wanted to know how the weather is down there in Puerto Rico, and to see if you found Adriana yet						
20	2004	YR	A	B	C	're in deep, deep trouble here. # -- # Cameron: How's the weather ? # Gabe: There's some minor turbulence and traffic, but nothing that's						
21	2003	ATWT	A	B	C	, this huge valley that's growing between us. Hey. How's the weather over there? Nice and warm over here. Why don't you come on						
22	2003	DAYS	A	B	C	? It would just be idtic small talk, like, " how's the weather where you are? Cause it's really hot here in salem. "						
23	2003	GL	A	B	C	. Bill: How you been, Dad? Bill: Mm-hmm. Billy: How's the weather ? Look, what, you and Michelle have some trouble? Bill: All right						
24	2003	GL	A	B	C	it's fine. We were just talking business. # Danny: How's the weather in New York? Salerno: Queens ain't New York, Danny, but it						

Fig. 7: List of examples in which how occurs in combination with the noun weather

If you compare the single sentences with each other, you will find a common sentence structure: *How's (or is/was) the weather (in ...)?*

Step 4: Searching for particular parts of speech – Finding adjectives that describe the weather

As a next step, you need to find adjectives that can be used in order to describe a specific weather situation. For this, you again need to use the “Collocates”-function (see step 3). In the lower box, click on “POS” again and this time select “adj.all” (meaning all kinds of adjectives). As the adjectives you are searching for can either occur directly in front of the noun *weather* (as in *xxx weather*) or up to three positions after it (as in *the weather is very xxx*), you need to click on number 1 on the left side and on number 3 on the right side. Your search box should then look like the one shown in Fig. 8:

Fig. 8: Collocate search box before clicking on “Find collocates”

However, you will probably get several different inflected versions of the same adjective (e.g. *warm* and *warmer*, *cold* and *colder*) if you leave your search box like this. If you want to know how frequent an adjective is regardless of its particular grammatical forms, you need to click on “Options”, which can be found in light grey letters in the right-hand corner, and select “Group by lemmas” (Fig. 9). A lemma is “a form of a word that appears as an entry in a dictionary and is used to represent all the other possible forms” (Cambridge Online Dictionary 2020).

The screenshot shows a search interface with the following elements:

- Navigation tabs: List, Chart, **Collocates**, Compare KWIC
- Search box: weather_nn* (Word/phrase), noun.ALL
- Collocates box: _j* (Collocates), adj.ALL
- Frequency scale: + 4 3 2 1 0 0 1 2 3 4 +
- Buttons: Find collocates, Reset
- Options menu: Sections, Texts/Virtual, Sort/Limit, **Options**
- Settings:
 - # HITS: 100
 - # KWIC: 200
 - GROUP BY: LEMMAS**
 - DISPLAY: RAW FREQ
 - SAVE LISTS: NO

Fig. 9: How to group words by lemmas

If you now click on “Find collocates”, a list similar to the one shown in Fig. 10 should appear on your screen.

HELP	<input type="checkbox"/>	CONTEXT	FREQ
1	<input type="checkbox"/>	[BAD]	87
2	<input type="checkbox"/>	[WARM]	41
3	<input type="checkbox"/>	[NICE]	31
4	<input type="checkbox"/>	[COLD]	26
5	<input type="checkbox"/>	[PERFECT]	20
6	<input type="checkbox"/>	[GOOD]	20
7	<input type="checkbox"/>	[GREAT]	20
8	<input type="checkbox"/>	[BEAUTIFUL]	16
9	<input type="checkbox"/>	[MAD]	9
10	<input type="checkbox"/>	[BETTER]	9
11	<input type="checkbox"/>	[NATIONAL]	8
12	<input type="checkbox"/>	[LOVELY]	7

Fig. 10: List of adjectives (grouped by lemmas) that frequently occur in combination with the noun *weather*

The list shown in Fig. 10 includes all adjectives that occur in combination with the noun *weather*, either directly in front of the word or up to two positions afterwards. To find out more about the specific context in which a certain adjective can be used in combination with *weather*, click on that particular adjective. Then, a new list will appear on your screen, showing all the example sentences in which this particular combination of the selected adjective and the noun *weather* occurs. If you click on the date on the far left, you can see more context and obtain information about the American soap opera from which a particular sentence was taken.

CLICK FOR MORE CONTEXT					<input type="checkbox"/>	[?]	SAVE LIST	CHOOSE LIST	CREATE NEW LIST	<input type="text"/>	[?]	SHOW DUPLICATES
1	2012	GH	A	B	C								
insisted on staying in that motel... why she took her newborn out in bad weather or wandered off into the woods. She never has to ask herself... if													
2	2011	AMC	A	B	C								
Hayward, Jack: Hayward's back in prison, isn't he? Jesse: Bad weather delayed the transport. He's still here downstairs, under guard. Erica: Then													
3	2011	AMC	A	B	C								
want to leave Lydia with a sitter. I mean, not in that bad weather and all. Because what if the power went out, and Lydia was so													
4	2011	GH	A	B	C								
the lodge? Dante: The 10 is, but 55 is the best in bad weather , and I'm assuming that's what the driver did. # Carly: Are													
5	2011	GH	A	B	C								
baggage cart out there. Sonny: Well, maybe he tripped because of the bad weather and he fell. Shawn: It must be convenient how people who get in your													
6	2010	AMC	A	B	C								
is doing. You know, Mom, he stood me up! Liza: Bad weather grounded his plane. Colby: Really? Liza: Yeah, really. Colby: I should													
7	2010	ATWT	A	B	C								
have got to be kidding me. (Door-closes) Luke: What's wrong? Reid: Bad weather in Dallas. Nothing commercial's flying in. Luke: Oh, good. That													
8	2010	ATWT	A	B	C								
I was living on the streets. I was living in the woods when the weather got bad . Stole what I could, foraged the rest, and one day													

Fig. 11: List of example sentences featuring the adjective *bad* in combination with the noun *weather*

Based on your findings, you can now choose some adjectives that you would like to introduce to your pupils in the upcoming lessons. When selecting these adjectives, keep in mind that they are sorted by frequency. Given that one of the core objectives of teaching English as a foreign language is the teaching of authentic English, it is advisable to only include adjectives in your lessons that are frequently used in combination with the noun *weather*.

Step 5: Categorizing your research results

As a last step, you need to categorize your results in order to use them for your lesson preparation. Table 1 serves as an example of what such a categorization could look like:

Table 1: How to describe the weather in different ways (neutral, positive, negative)

Describing the weather in a neutral way	Describing the weather in a subjective way (positive)	Describing the weather in a subjective way (negative)
The weather is ...	The weather is ...	The weather is ...
warm	nice	bad
cold	perfect	terrible
hot	good	horrible
stormy	great	awful
windy	beautiful	
sunny	lovely	
chilly	fine	

3.1.2 Task creation – Designing corpus-based teaching materials

3.1.2.1 Warm-up activity

Step 6: Creating a warm-up activity in the form of a dialogue

As an **introduction** and **warm-up activity**, you can create a fictional dialogue between yourself as the teacher and a fictional person or the class mascot, talking about your plans for the holidays. At this point, it has to be noted that the pupils should already be familiar with typical vocabulary, phrases and sentence structures that can be used to greet somebody. Ideally, information regarding these phrases, sentence structures and vocabulary has also been derived from a corpus in previous lesson preparations. The warm-up activity thus serves both as a revision of old knowledge and as an introduction to new skills and vocabulary. The introductory dialogue could look like this:

Teacher: Hey (name of class mascot)! How are you?

Mascot: Hello Mr./Mrs. (name of teacher). I'm good, thanks. How are you?

Teacher: I'm good, too. What are you going to do in your holidays?

Mascot: I'm going to (country/city).

Teacher: Wow, you are going to (country/city)? That sounds like a lot of fun! How's the weather in (country/city)?

Mascot: It's chilly in (country/city).

Teacher: Brrrr, I don't like cold weather.

Mascot: What are you going to do in your holidays then?

Teacher: I'm going to (country/city).

Mascot: How's the weather in (country/city) then?

Teacher: It's nice and warm so I can go to the beach!

Mascot: That sounds like a great plan!

Alternatively, you can also show an appropriate video in which people are talking about the weather and afterwards ask your pupils what the video was about. However, it can be quite difficult and time-consuming to find a video that is suitable for this particular warm-up activity. Therefore, it is advisable to create such a fictional dialogue.

3.1.2.2 First task

Step 7: Creating word and picture cards

The warm-up activity provides the transition to the **first task**, which introduces new vocabulary. To create the material for this task, create picture cards that correspond to the adjectives listed in Table 1. All the cliparts used in this chapter are from <https://publicdomainvectors.org> and have the Common Creative license CC0.

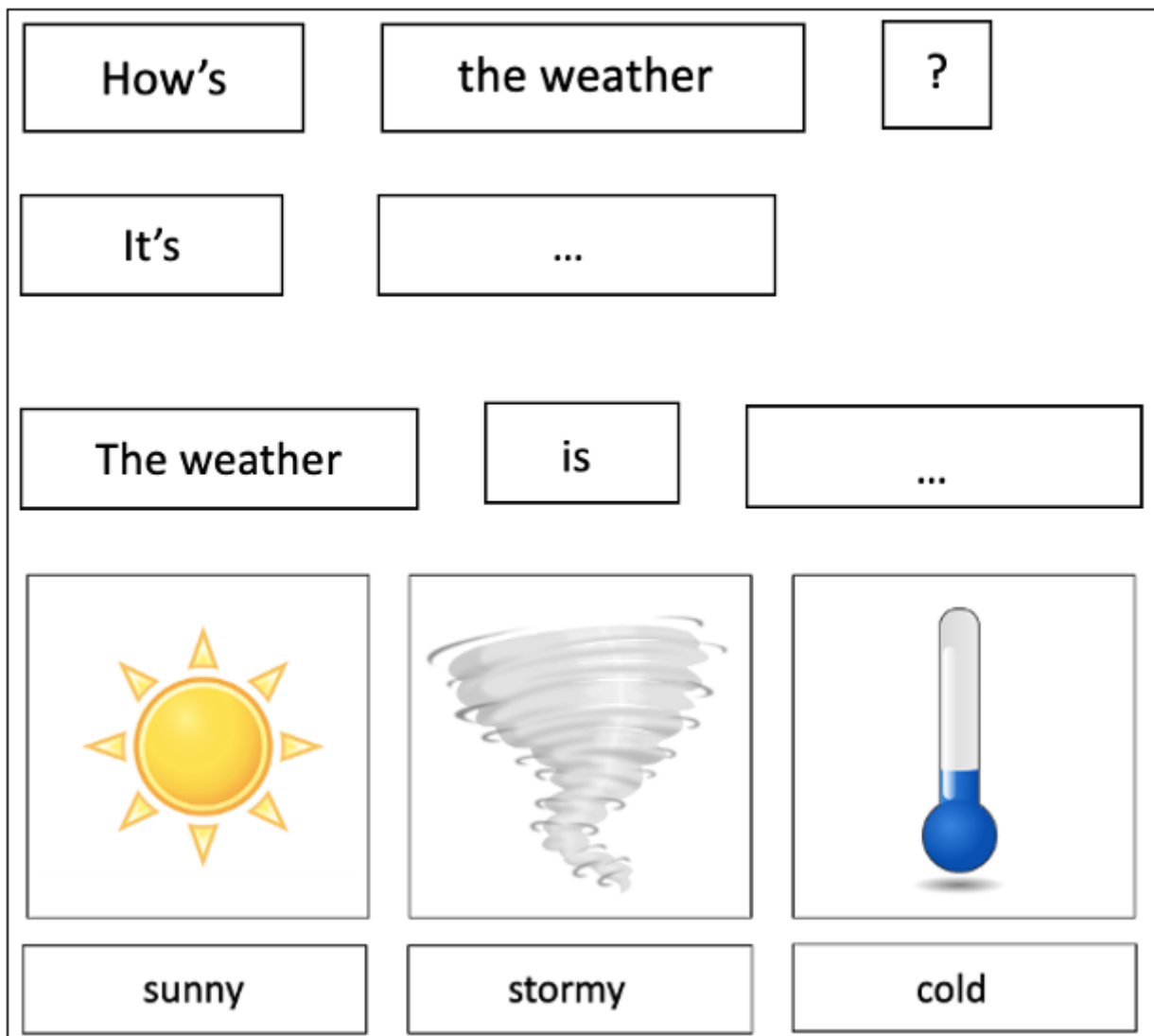


Fig. 12: Example word and picture cards (clip arts: <https://publicdomainvectors.org>)

3.1.2.3 Second task

Step 8: Designing a dice template

The **second task** aims to motivate the pupils to actually use the vocabulary in a dialogue or other forms of communication with the help of a weather dice featuring the clip arts you already used in order to design your picture cards. You can either create the ready-to-use dice on your own or provide your pupils with a template they can cut out and glue together. Such a template could look like this:

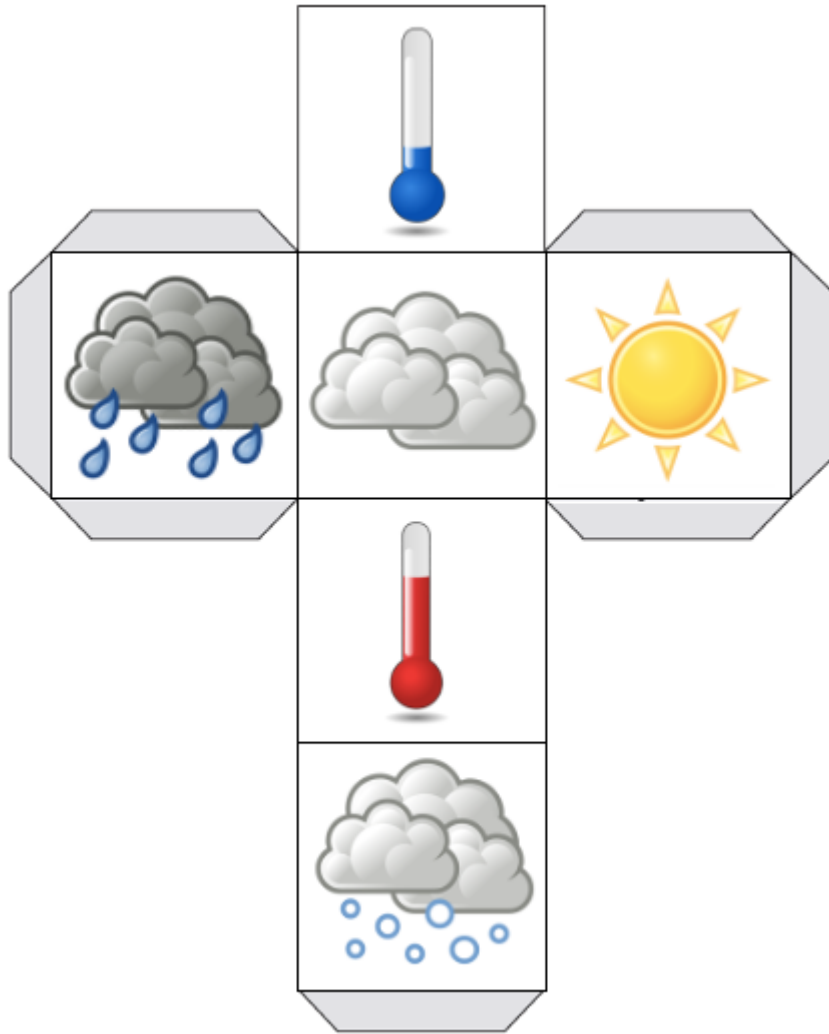


Fig. 13: Example dice template (clip arts: <https://publicdomainvectors.org>)

3.1.2.4 Third task

Step 9: Selecting a weather report video and designing play cards

The **third task** can be introduced by a weather report. Try to find a short weather report that can be easily understood by beginners of the English language and that, at the same time, can also be fun and motivating for primary school pupils. A great example is the following video by Maple Leaf Learning: <https://www.youtube.com/watch?v=gpBuaU5OPi8>.

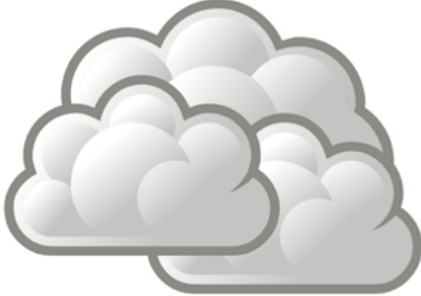
Based on this video, you can create play cards that illustrate the weather in different cities/countries and include a question to ask. You can use the same clip arts used in the picture cards pinned to the board during the first session.

Barcelona
In Barcelona, the weather is ...



I want to travel to Paris.
How is the weather in Paris?

Paris
In Paris, the weather is ...



I want to travel to London.
How is the weather in London?

Fig. 14: Example play cards (clip arts: <https://publicdomainvectors.org>)

3.2 Lesson sequence

Table 2: Lesson sequence “How’s the weather?”

Lesson	Task	Implementation	Interaction form	Medium/Material
1	Warm-up activity	Revision of previously acquired vocabulary, introduction of new vocabulary	Plenum	Fictional dialogue (or video), class mascot
	First task: Asking about the weather	The pupils work out the sentence structure of <i>How's the weather (in...)?</i> and the corresponding answer. They familiarize themselves with this question and its possible answers.	Plenum (possibly sitting in a circle)	Word cards, picture cards
2	Second task: Consolidating and deepening the pupils' knowledge	In pairs, pupils make use of the weather dice. One pupil asks the question <i>How's the weather?</i> The other throws the dice and gives the corresponding answer.	Pairs	Weather dice/dice templates
3	Third task: Creating a classroom conversation about the weather	All class members work together and ask each other how the weather is in a certain city/country they want to travel to. The different play cards appoint one city/country and the corresponding information to a pupil.	Plenum	Weather report video, play cards
*	Additional task: Creating a weather report	With the help of some pre-fabricated sentence structures, the pupils create their own fictional weather reports.	Pairs or small groups	Worksheet with sentence structures/chunks, camera/tablet
*	Additional task: Working on a weather diary	After the first lessons, the pupils can start collecting daily information about the weather and record their observations in the form of a diary.	Individual work	Weather diary with pages for each day of the remaining unit and a page with supporting sentence structures/chunks

4 Options and further ideas

Having completed the first steps of the unit presented in this chapter, pupils could also be encouraged to further develop or deepen their acquired skills by creating their own weather report or weather diary. The report could be produced in groups. The word and picture cards introduced beforehand, as well as cards with supporting sentence structures, may serve as additional support for weaker pupils. The sentence

structures on the support cards may be based on the weather report video, which will already be familiar to the pupils from task 3.

The weather diary could also be part of the pupils' homework. The pages for the first few days may contain sentences with gaps that simply need to be filled in. As pupils progress, the design of the pages could then become more independent in writing their weather diary entries. By the end of the diary, pupils should be able to fill a page entirely on their own. Entries could include general information about the current weather situation, as well as information regarding the temperature or the next day's weather forecast.

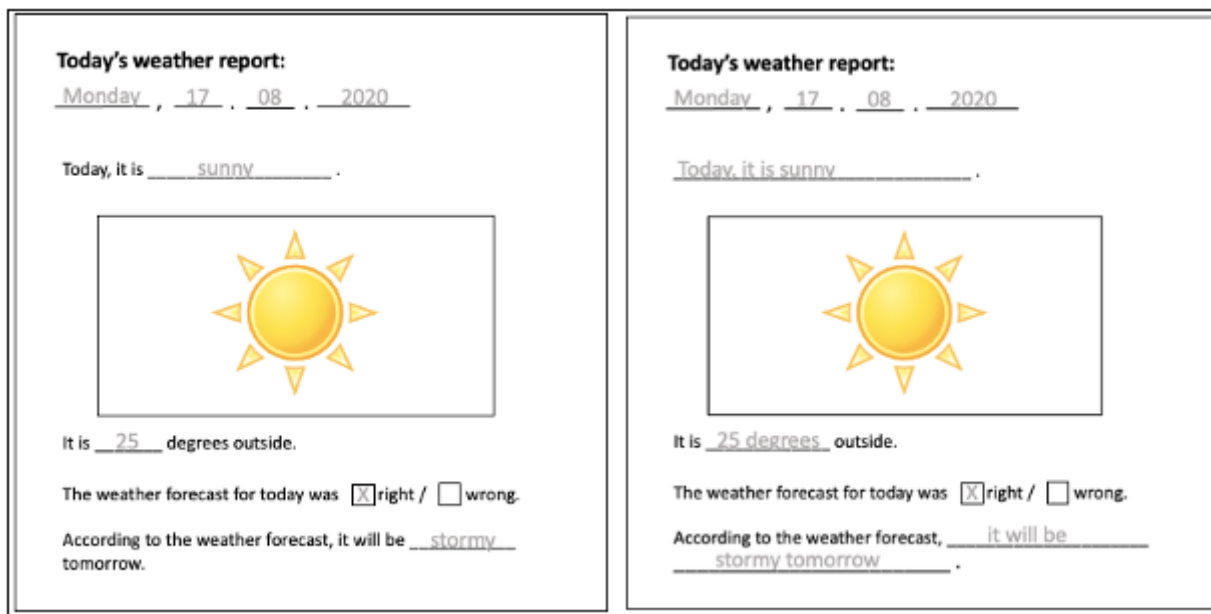


Fig. 15: Example pages for the beginning of the weather diary task (clip arts: <https://publicdomainvectors.org>)

5 Caveats and limitations

[Click here to continue with your search](#)

[简体中文](#) [中文](#) [한국어](#) [Русский](#) [العربية](#)

There is no cost for basic access to English-Corpora.org. But you will occasionally see this message, which asks you to consider upgrading to a "premium" account.

if you have a **premium account** (\$10.00 per month, \$30.00 for one year, \$55.00 for two years, or \$75.00 for three years), then you will not see this message anymore. You will also have increased access to the corpora, more features, and you will help to support English-Corpora.org.

if you are a student, your school or university can also get an **academic license**, and all of you will have the benefits of a premium account.

Thanks for your support.

Fig. 16: How to continue your research without creating a paid premium account

When using English-Corpora.org to derive data for the purpose of designing teaching materials, some difficulties might be encountered. Firstly, a free account on this platform only allows the user to conduct a limited number of searches per day. In addition, after having used the searching tools for a while, a notification informing you about the possibility to upgrade your account will be displayed (Fig. 16).

If this notification appears on your screen, you simply need to wait for a few seconds. Then, a new button appears, which enables you to continue your search without creating a premium account.

Secondly, there are three ways of obtaining the clip arts that you will need for the tasks described in this chapter. You can either create them yourself, buy them from an illustrator or make use of open educational resources. In the latter two cases, it is mandatory that you are aware of and follow the respective copyright guidelines of the clip arts that you use. If you buy your clip arts from an illustrator, you are often only allowed to use them for your own personal crafting and classroom use. Thus, you are generally neither allowed to share these clip arts nor use them commercially. Consequently, if you intend to publish and resell your teaching materials, it is advisable to make use of open, publicly accessible educational resources. They are openly licensed, which means that users are allowed to reuse and even publish these materials and resources in their own works. The clip arts included in this chapter are open educational resources, too, and are covered by the Common Creative license CC0.

Lastly, it should be noted that there are different ways of how to ask somebody about the weather. It is also possible, for example, to ask *What is the weather like today?* Depending on which corpus you use and which kind of language data this corpus is based on (British or American English, spoken or written English, etc.) your data collection results may well differ from the ones presented in this chapter (see section 3.1.1). We deliberately decided to only introduce one way of asking about the weather in our proposed tasks because there is a high probability that pupils will otherwise mix several variants, thereby creating unidiomatic questions, e.g., **How's the weather like today?* Nevertheless, it is, of course, possible to set the same tasks using the question *What is the weather like today?* or any other idiomatic expression.

6 Conclusion

This chapter has shown that working with corpora enables teachers to design innovative and authentic teaching materials and thus allows pupils to acquire useful, idiomatic phrases for conversations in English. Though the data compilation process may initially seem quite time-consuming and more complex than working with a conventional textbook, working with corpora can compensate for some of the weaknesses of the materials traditionally used in the primary school classroom (see section 1). However, in order to take advantage of these benefits, it is important to consider the appropriateness of the corpus chosen. Once collected, the corpus data can be easily stored and reused numerous times. Finally, the processes of collecting corpus data and developing motivating teaching materials on the basis of this data can be applied to any other topic of the EFL classroom that aims to enable (primary school) pupils to acquire communicative competences based on authentic English.

7 Resources and references

Corpus of American Soap Operas. <https://www.english-corpora.org/soap/> (25 February, 2021).

Maple Leaf Learning. 2011. International Weather Report – Simple Skits. <https://www.youtube.com/watch?v=gpBuaU5OPi8> (25 February, 2021).

Public Domain Vectors. <https://publicdomainvectors.org> (25 February, 2021).

Cambridge Online Dictionary. 2020. Lemma. <https://dictionary.cambridge.org/de/worterbuch/englisch/lemma> (04 August, 2020).

Jones, Christian & Tania Horak. 2014. Leave it out! The use of soap operas as models of spoken discourse in the ELT classroom. *Journal of Language Teaching and Learning* 4(1). 1-14.

Niedersächsisches Kultusministerium. 2006. Kerncurriculum für die Grundschule, Schuljahrgänge 3-4: Englisch. <https://docplayer.org/11297531-Niedersaechsisches-kultusministerium-kerncurriculum-fuer-die-grundschule-schuljahrgaenge-3-4-englischnieder-sachsen.html> (25 February, 2021).

O'Keeffe, Anne, Michael McCarthy & Ronald Carter. 2007. *From corpus to classroom language use and language teaching*. Cambridge: Cambridge University Press.

Quaglio, Paulo. 2009. *Television dialogue: The sitcom Friends vs natural conversation*. Amsterdam; Philadelphia: John Benjamins.

Surkamp, Carola & Britta Viebrock (eds.). 2018. *Teaching English as a foreign language: An introduction*. Stuttgart: J.B. Metzler.

Thaler, Engelbert. 2012. *Englisch unterrichten: Grundlagen, Kompetenzen, Methoden*. Berlin: Cornelsen.

2. "How do I get there?": Travelling by various means of transport

MAGDALENA FUCHS AND ANNIKA LIEBELT

1 Introduction and rationale

Your turn: Choose five words from the grid and use them to write a short text about your last holiday or trip. (Green Line 4 G9, 2017: 9)

This example instruction, taken from a secondary school English coursebook widely used in Germany, illustrates that textbooks often focus on the learning of individual words, while the teaching of idiomatic phrases frequently comes short. The proposed lesson follows a different approach to language learning, in which the learner learns to think in linguistic chunks. It provides language learners with the necessary scaffolding to successfully go from learning new words to using them in context.

This chapter provides teachers with tasks that are close to students' reality and are connected to the newly learned expressions. In accordance with the overall aim of communicative language teaching, the students discuss travel plans in the form of text messages, explore strong collocations in written and spoken language and analyse the structures of a selection of phrases. As intended by the curriculum of Lower Saxony (Germany), students learn how to give information and make plans in English in an interactive context (Niedersächsisches Kultusministerium 2015: 17-18) with the help of chunks and common expressions that they gather from authentic, yet simple language. The focus lies on verb-noun and prepositional collocates.

The proposed lesson aims to support students in becoming autonomous language learners by means of scaffolding. When working with younger students and corpora, there is a fine line between student-corpus interaction and information overload, so the research tasks are kept simple (cf. Friginal 2018: 168). According to the local curriculum, students are to learn how to use various writing resources and tools in order to become increasingly independent (Niedersächsisches Kultusministerium 2015: 20).

For teachers, this chapter provides a step-by-step guide on how to find authentic chunks, questions, and example sentences for communicative travel situations, how to visualise and use them in class. The following corpus tasks encourage students to think beyond the labels of "correct" and "incorrect" language use, and to explore the special features of spoken English.

Outline and objectives

Topic: *Travelling by various means of transport*

Teaching aim: show how authentic speakers' expressions can be used to create context-bound activities with relevance for real life. Help learners familiarise themselves with the corpus tool *SkeLL* (Sketch Engine for Language Learners) and encourage them to make use of the advantages that corpus tools offer.

Duration: 60-minute class. Single activities can be integrated into an individual lesson plan, optional tasks are included for an extended class.

Targeted age group: 7th or 8th grade

Preparation time: approx. 30 mins, less if already familiar with the corpus tool.

Learning objectives:

- Learners are able to identify commonly used verb collocations for means of transportation by using the Word Sketch function on SkeLL.
- Learners are able to reply to people asking for directions (e.g., "How do I get from the airport to your house?").
- Learners are able to explain to an exchange student how to travel to their destination in Germany in the form of a written dialogue. This enables students to use various expressions encountered in the corpus.

2 Corpus, tools and methods

Content-wise, the lesson follows the idea of learning to use collocates in context. On a methodological level, its aim is to help learners become proficient users of the corpus tool SkeLL. SkeLL is a free corpus tool specifically designed for language learners created by Sketch Engine. It offers a very intuitive user-friendly interface, which makes it highly suitable for learner-corpus interaction. Another positive aspect is that the tool functions well on smartphones so that it is not even necessary to set up laptops in class.

From intermediate proficiency level onwards, secondary school students should be encouraged to reflect on their language use more autonomously in order to become independent language learners. In the proposed lesson, students get the opportunity to work with a corpus tool and to do research themselves, which encourages language awareness and promises to be highly motivating.

The first function used for the analysis is the "Word sketch" function. It provides users with lists of words surrounding the search word, arranged in categories (collocating verbs, nouns, adjectives etc.). This means that it is possible to instantly see the most typically used verbs which have the search word as object. The corpus tool also offers example sentences including the search word, which provides learners with information on how the word appears in context.

The corpora used for the lesson are the ones preloaded on SkeLL: they "contain a varied collection of texts such as news, academic papers, Wikipedia articles, open-source books, web pages, discussion forums, blogs, etc. to provide a good example of how language is used in everyday, standard, formal and professional context" (www.sketchengine.eu/SkeLL/).

In addition, this chapter provides step-by-step instructions to creating a worksheet with Sketch Engine, which is a commercial platform and therefore only accessible with an account beyond the 30-day trial period, but which offers a wider choice of corpora and corpus functions that may prove very useful for teachers.

3 Step-by-step guide

3.1 Teacher preparation

It is advisable to give students an introduction on corpora, the SkeLL tool and its functions in advance. The teacher and each student (or student pair) need a device connected to the internet. A laptop would be ideal, but SkeLL also works on mobile devices. This corpus platform is not used for the warm-up or the pre-task.

3.1.1 Collecting expressions for worksheet 1

Open SkeLL in your internet browser (<https://skell.sketchengine.eu/#home>).



Fig. 1: SkeLL search



Fig. 2: Search options

Make sure to select “English” next to the search bar (Fig. 1). Type in a keyword, for example *bus*. Click on “bus in Word Sketch” in the appearing drop-down menu or select “Word sketch” after having started the search (Fig. 2). The word sketch then provides you with a well-structured overview of the strongest collocations of this keyword.

Select “noun” in the task bar to search only for phrases with *bus* as the noun (Fig. 3). This is particularly important for words like *train* that are also verbs.



Fig. 3: Settings

For task 1, you want to find verbs that are strongly associated with *bus* as an object, so focus on that column first (Fig. 4). It displays a variety of verbs you can choose from. Caution: not all verb listed will be relevant to your task. It is the teacher’s responsibility to explain to the class why some collocates might not fit the task: e.g., if we are looking for verb-noun collocations related to the topic of travelling, *articulate*, *park* and *drive*, all of which are also among the strongest collocates of *bus*, are not relevant.

verbs with bus as object	
1. <u>board</u>	boarded a bus
2. <u>ride</u>	ride the bus
3. <u>articulate</u>	articulated buses
4. <u>park</u>	park the bus
5. <u>catch</u>	catch a bus
6. <u>operate</u>	buses operated
7. <u>drive</u>	driving the bus
8. <u>equip</u>	buses are equipped
9. <u>exit</u>	exit the bus

Fig. 4: Column “verbs with ‘bus’ as object”

To see example phrases that feature one of these collocations, click on the collocate (e.g., *catch*) (Fig. 5).

catch + bus 0.32 hits per million

1. Had set off to **catch** the **bus** .
2. A long walk, then **catch** a **bus** .
3. Suppose Homer wants to **catch** a stationary **bus** .
4. What remains difficult in the country is **catching** the **bus** .
5. I could **catch** a **bus** back into town.

Fig. 5: Concordance lines with catch and bus

You can collect phrases suitable for your students for worksheet 2 on SkeLL (cf. section 4.1.1). The example phrases in example worksheet 2 were, however, taken from the Web English 2018 corpus using the Sketch Engine platform. Because of its considerably larger range of corpora and options, teachers will find many more examples there, so they can more easily vary the level of tasks. However, its interface is not as user-friendly as SkeLL, so this tool may be more suitable for research purposes and teacher preparation than for student-corpus interaction. Below, you will find a step-by-step guide for teachers for finding good examples on Sketch Engine.

3.1.2 Finding phrases for Worksheet 2 with Sketch Engine

3.1.2.1 Registration and login on Sketch Engine

Create an account on Sketch Engine or log in with your institutional account (Fig. 6). Participating EU institutions can access the tool for free until 2022 through the ELEXIS project. Note that, unless your school is taking part in this project, Sketch Engine is not free to use. There is, however, a 30-day free trial option.

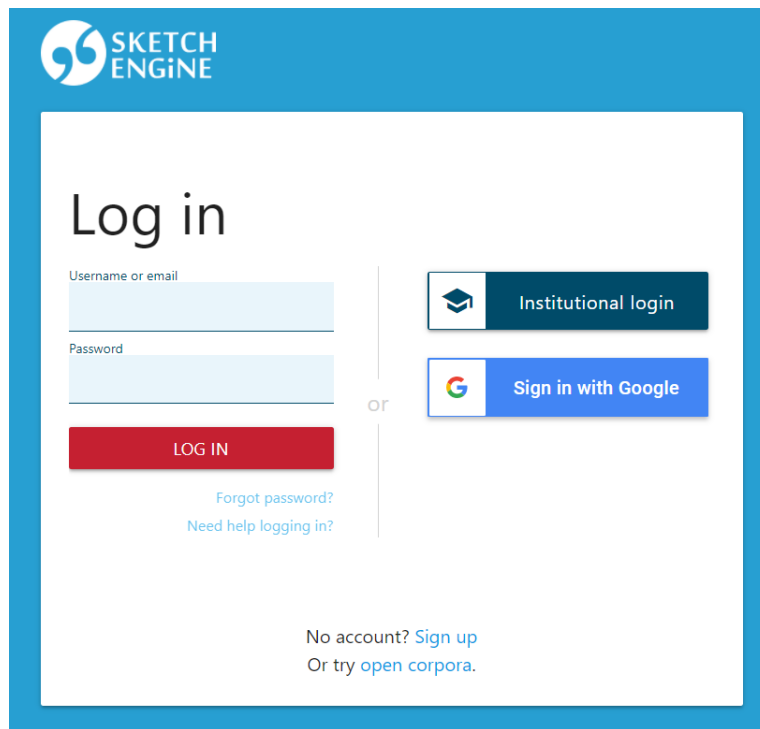


Fig. 6: Sketch Engine log in screen

The Sketch Engine dashboard provides you with an overview of all the platform's key functions.

3.1.2.2 Selecting a suitable corpus and using the “Word sketch” function

In the search bar at the top of the dashboard view, select the English Web 2018 corpus (enTenTen18) (Fig. 7). With over 21 billion words, it is considerably bigger than other corpora. Note that the texts contained in the corpus are not only produced by native English speakers. Considering the size of the corpus, though, this is unlikely to have an impact on the search results.

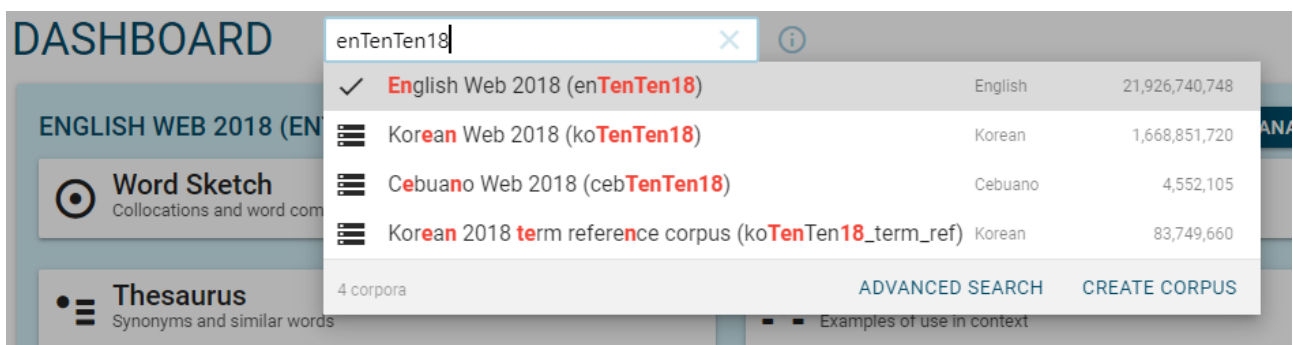


Fig. 7: Corpus selection

Just like SkeLL, Sketch Engine can be used to create word sketches. Select this function on the left side of the dashboard view. Then, select the “Advanced search” tab in the top bar. Type a noun referring to a means of transportation (e.g., *bus*) into the search bar. Sketch Engine will search for *bus* as a lemma, meaning that it will search for both singular and plural forms of the word. Select “noun” as the part of speech below the search bar. Click “GO” to start the search (Fig. 8).

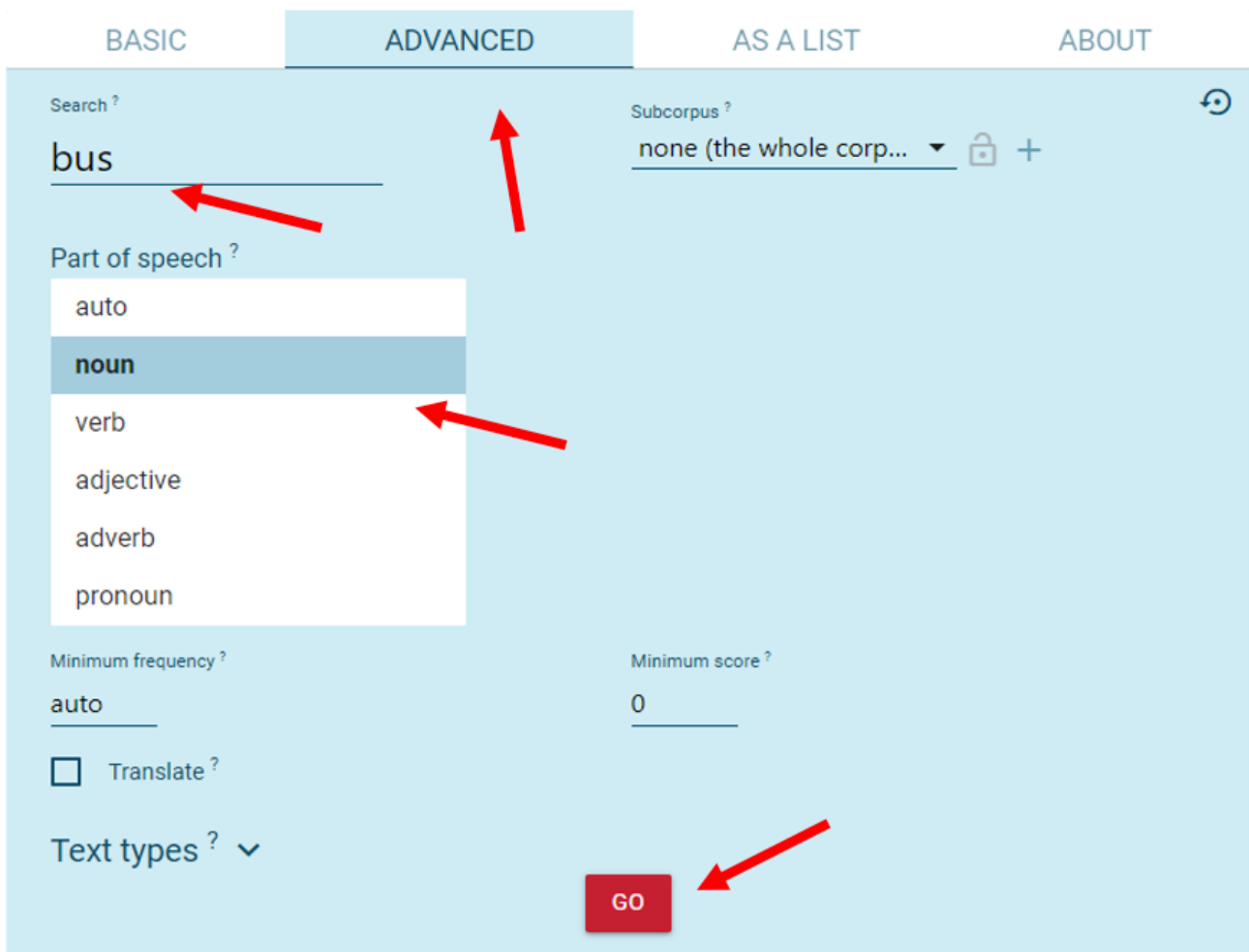


Fig. 8: Word sketch and advanced search

The symbols in the top-right corner help you organise your results. Click on the eye symbol to sort the collocations according to their collocational strength if you want the results to look like they do on SkeLL, which always displays the strongest, most typical collocates of a certain keyword. If you sort by frequency, many of the collocates listed in the word sketch will probably be words which occur with a high frequency in the language in general (e.g., *take* or *get*). Choose the option that meets your requirements best or, for a good mixture, you can even use both (Fig. 9).

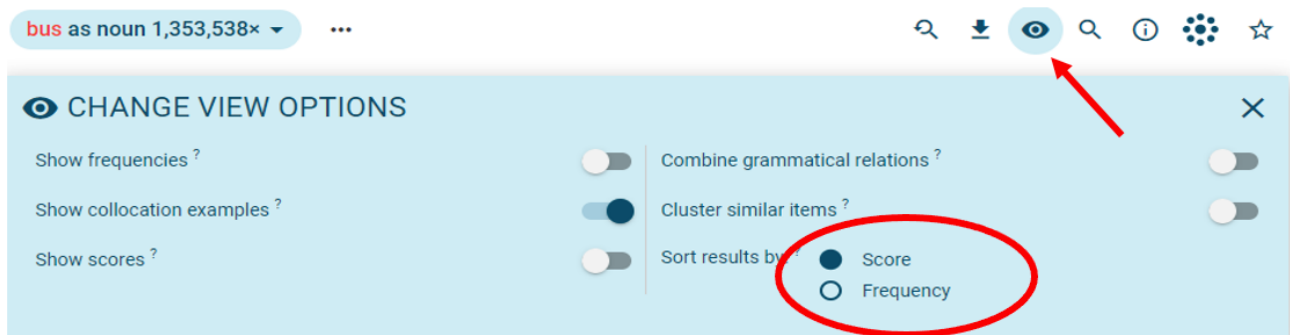


Fig. 9: View options for the word sketch

3.1.2.3 Exploring concordances

Again, focus on the column “verbs with ‘bus’ as object”. To see how the verb-noun collocations are used in context, click on “...” next to a collocate and open the concordance lines for “bus + catch” in a new tab for a better overview (Fig. 10).

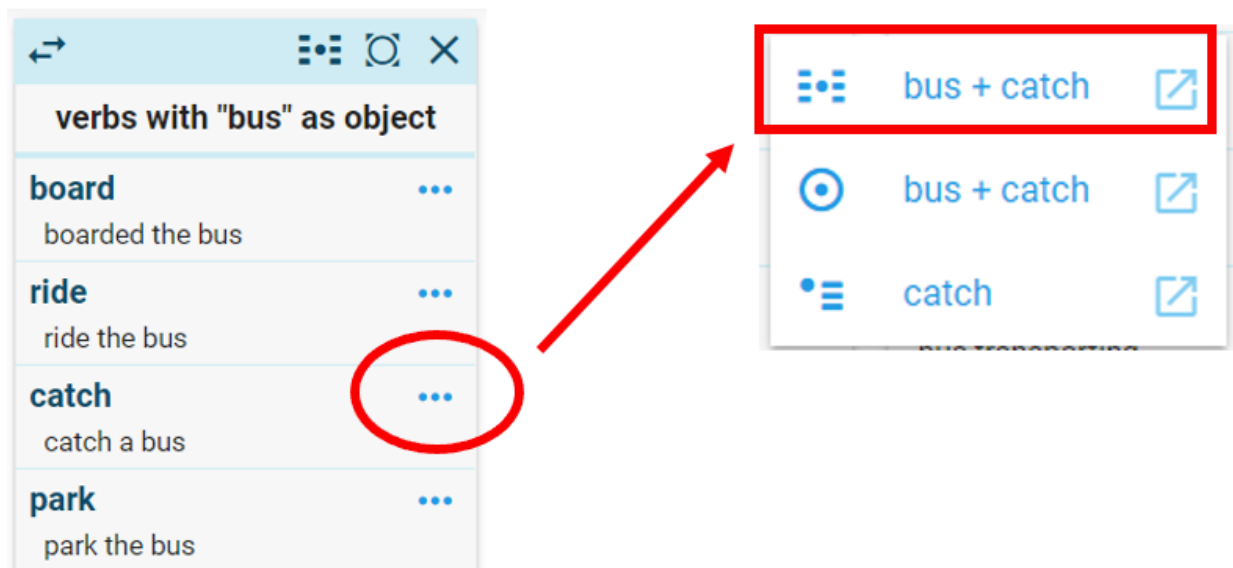


Fig. 10: Column “verbs with ‘bus’ as object” and further options

It is advisable to select a random sample of 100 or 200. Click on the question mark symbol in the toolbar above, then click “GO” (Fig. 11).

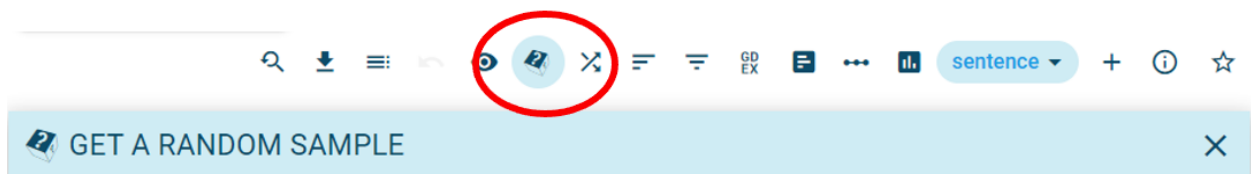


Fig. 11: Random sample

Select “KWIC” to see the collocation “bus + catch” used in a larger context, with the keyword *bus* displayed in the centre (Fig. 12). If you prefer to see the collocation used in complete sentences, select the option “Sentence” (Fig. 13).

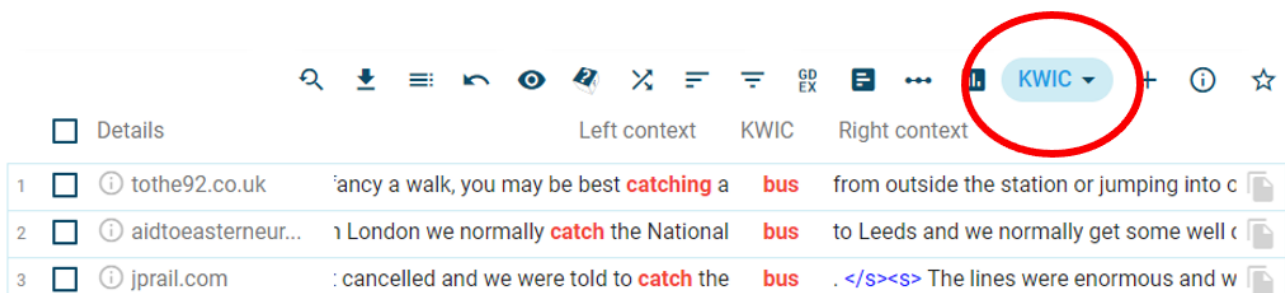


Fig. 12: The “KWIC” function

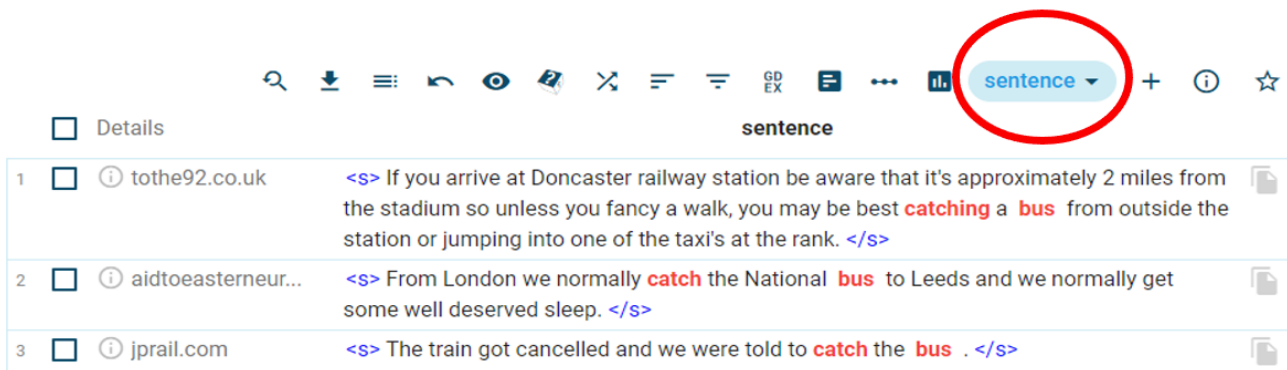


Fig. 13: The "Sentence" function

Click on "bus" to see the extended context of the concordance line. You can see an even larger excerpt by clicking on "..." at the top or at the bottom of the of the paragraph (Fig. 14). Pick single sentences or short dialogue excerpts you want to include in your worksheet.

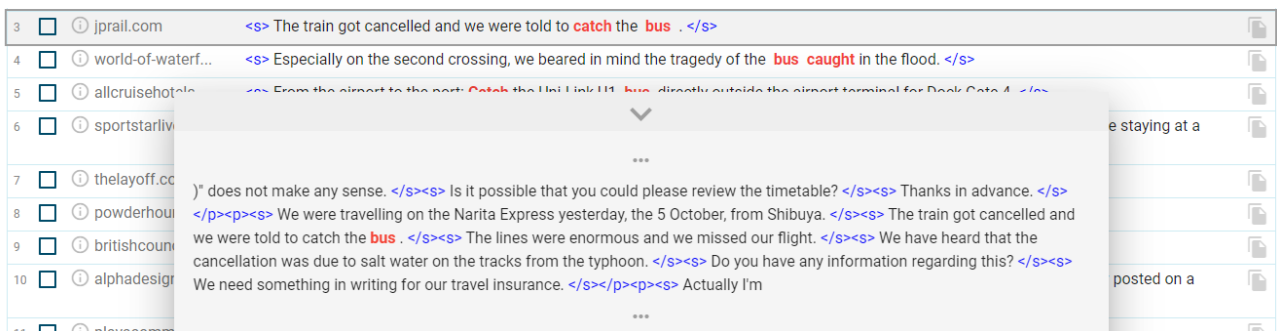


Fig. 14: Extended context of concordances

3.1.2.4 Finding "good dictionary examples"

This function helps you to filter the concordances to find shorter, simpler example sentences. Click on the "Good dictionary examples" symbol in the toolbar, then click "GO" (Fig. 15). You can always undo certain steps of your search or go back and forth between your steps. Click on the "x" displayed in the search history above to undo a step (Fig. 16).

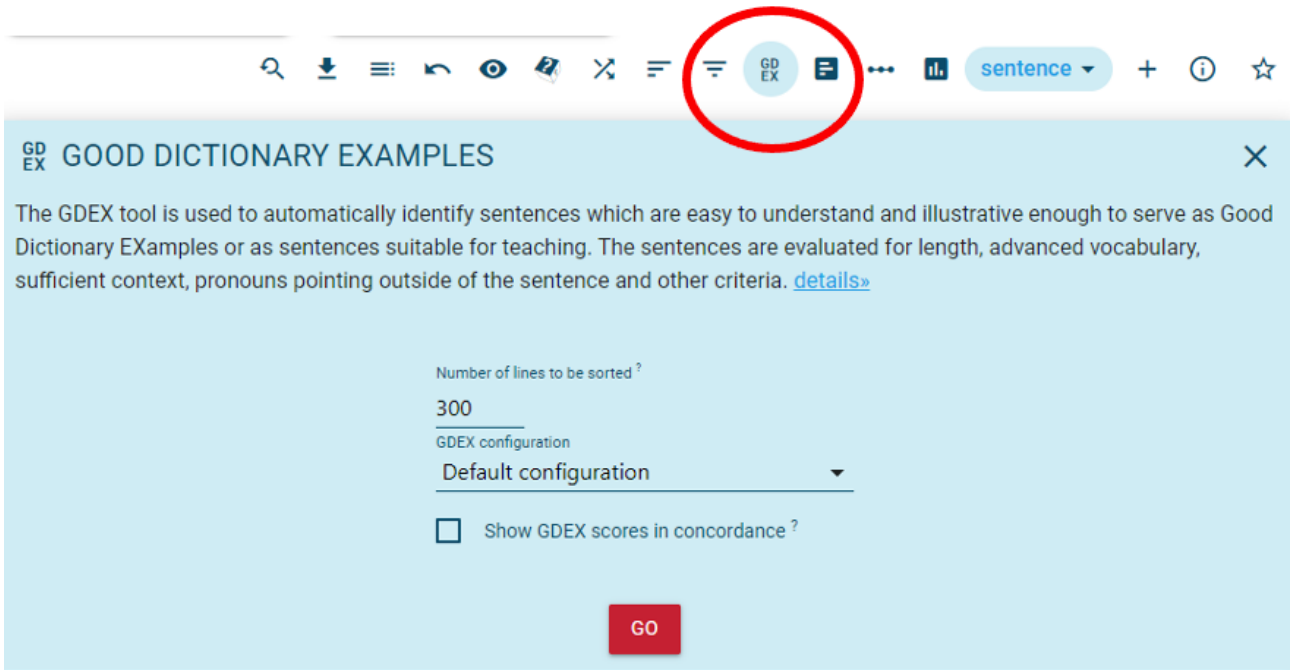


Fig. 15: Settings for “good dictionary examples”

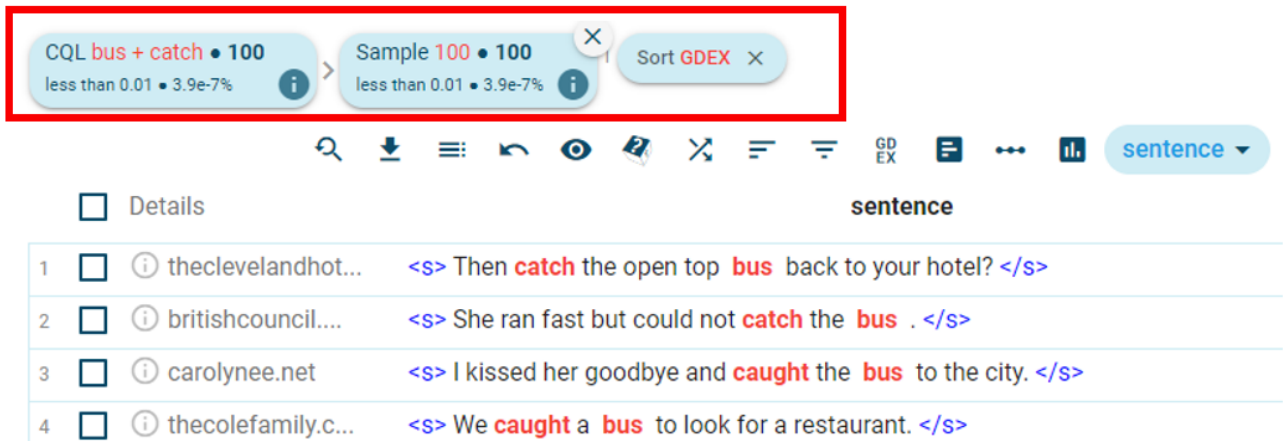


Fig. 16: “Good dictionary examples”

4 Lesson plan

Table 1: Lesson plan “How do I get there?”

Phase	Learning activity	Instructions	Interaction form/ materials
Warm-up 10mins	Students tell their neighbour how they travelled to a place they visited.	Ask students to tell their neighbour about a trip they went on. Where did they go and how did they get there?	In pairs
Pre-task 5 mins	Learners name means of transport.	Note down students' ideas in a prepared mind map and then display possible verb collocates.	Projector/ PowerPoint Plenum
Corpus research 20 mins	Students create word sketches for different means of transport to find out which verbs work with which nouns (as object). Students write down example sentences from the concordance lines.	Explain instructions on worksheet 1 and present an exemplary search on SkeLL on the whiteboard. Support students if there are any questions.	Worksheet 1 In pairs
Collecting results 10	A few students present in class how they created the word sketch.	Complete the mind map with the verbs the students found in the word sketches.	Plenum
Further analysis 10	Students analyse corpus examples of prepositional phrases featuring the noun.	Hand out worksheet 2 and explains instructions.	Worksheet 2 Single + in pairs
Writing a dialogue 15 mins	Students give directions by completing a chat dialogue.	Hand out worksheet 3.	Worksheet 3 In pairs
Presentation 10 min	Students volunteer to read out their dialogues in class.	Ask students to present their dialogues in class.	Plenum

4.1 Instructions

Begin the lesson with a short warm-up activity. Students tell their neighbours how they travelled to a place they have visited before. Ask students to name means of transport and show them verbs that might appear in combination with the nouns. In the following work phase, the class conducts their own research using SkeLL to find out which verbs and nouns are commonly used as collocations in English for giving directions.

Before the class starts to work independently in pairs, present an exemplary search on SkeLL by sharing your screen and projecting it onto the wall or the whiteboard. It is important to highlight that the students should focus on the *context* of the collocations, i.e., the example sentences in which the search words appear. Worksheet 1 provides the students with a step-by-step guide on how to conduct the research and tells them which words to search for.

On the basis of the corpus tasks, the students transfer their findings to a written dialogue in which they give directions to an exchange student (see worksheet 3). By writing out the dialogue rather than giving directions verbally, learners cannot include gestures and point at directions or things, but are forced to use their (newly acquired) vocabulary to communicate. Nevertheless, if there is time left, let single pairs of students read out – or even act out – their chat messages.

4.1.1 Tasks

Worksheet 1: Doing your own corpus research



Follow the instructions to create word sketches. Then complete the table.

Which of the following verbs appear with each search word?

call | board | ride | catch | rent | drive | park | take | hail

Search word	3 verbs (see above) with "XY" as object	Click on a verb and write down an example sentence from the <i>Examples</i> .
Example: <i>car</i>	<i>drive, park, rent</i>	<i>He loved <u>driving</u> expensive, fast cars.</i>
bus		
train		
taxi / cab		
bike		

Instructions:

- Open SkeLL.
- Select the "Word sketch" function.
- Type a word for a means of transport into the search bar and select "noun", then press "enter".
- Look at the column "verbs with 'bus' as object"
- Fill in the worksheet – find three verbs per means of transportation.
- Open the example sentences in a new tab and find suitable examples.
- If you want to search for another word for a means of transport, first go back to the word sketch. Then type in a new word into the search bar.

Fig. 17: Worksheet 1 – Doing your own corpus research

Worksheet 2: Analysing sentence structures



A corpus shows you words and phrases from many different situations. Take a look at these example sentences from the English Web 2018 corpus. **Highlight** information you get on **the starting point**, **the destination**, and **the time**.



Do you notice a pattern? Describe the pattern. Work with a partner.



For corpus experts: You can work with these concordances **or** with SkeLL again. Go back to the examples and look for expressions describing the starting point, the destination, and the time. Create a word cloud.

1	Example: Next you have to take a	bus	from Riga to Ainazi in the Central Bus station.
2	Take that money and catch the first	train	for Chicago.
3	We [...] enjoyed the show & then caught a	bus	back to the city centre at 10.45pm.
4	I could easily catch the 6am	train	the next morning.
5	First you will need to get the	bus	from either [station 1] or [station 2] to Wareham station.
6	Very early in the morning, I hailed a	taxi	and made my way to Ottawa International Airport to catch my flight.
7	I thought we'd take the	train	into the city late this afternoon.
8	I even rode a	bike	to work today.
9	Yet another option: take a	bus	from the train station (Santa Maria Novella) to the airport.
10	I suggest you call a	cab	to take you home.
11	At about three p.m. we boarded our	train	to Fairbanks.

Fig. 18: Worksheet 2 – Analysing sentence structures



Worksheet 3: How do I get to your house?

Alex is an exchange student from London, who is coming to stay with you. Help him find the way from the airport to your house. Use the collocations you found on SKELL.



In your notebook, answer Alex's questions. Use the information from the speech bubbles.

ALEX: Where do I have to go when I arrive at the airport?



Arrival at Hanover airport: 9:40 am |



→ shuttle bus, number 27

ALEX: Great! What do I do when I've found the bus stop?



Departure at Hanover airport: 10:05 am

→ Arrival at Hanover central station: 10:23 am

ALEX: All right! What time is the train to Osnabrück?



Platform 12, dep. 10:37 am

→ Arrival at Osnabrück central station: 11:54 am

ALEX: Oh! That sounds doable. How do I get to your house from the station?



(Taxi)

OR



OR



pick you up



ALEX: Ok, thanks! I think I'll find the way. I'll text you if I have a problem. See you very soon!

...



Fig. 19: Worksheet 3 – "How do I get to your house?"

4.1.2 Follow-up activities

Another idea is to plan a real class trip together with the group. The students could create a flyer or a travel guide in which they give directions and conduct more research on the topic of travelling on SkeLL. Let students research what other aspects might be relevant when planning a trip and create more word sketches related to the topic. If there is a bit more time, it might also be interesting to search online for authentic public transport connections and practise giving corresponding directions. This would encourage students to focus on relevant information and to filter them according to their needs.

The proposed corpus activities might also serve as an introduction to the topic of travelling: following the ideas presented in this chapter, students could use the corpus tool to collect information on various destinations (cities, countries, sights, etc.), learn new adjectives to describe a trip or an event and use their knowledge to write a travel report, blog post or a diary entry.

5 Options and further ideas

The corpus tasks designed for this lesson allow for differentiation: teachers can decide how much scaffolding is needed, depending on the group of learners. For instance, the list of verb collocations might be prepared by the teacher or else be collected and developed by the learners themselves. Moreover, it is possible to let students find their own example sentences in SkeLL for further analysis of concordance lines. For the production task, some students can be challenged to create the whole dialogue themselves.

For a more realistic setting, it is possible to use exchange platforms like Erasmus eTwinning so that students can communicate with other English learners across Europe, or let students chat with their classmates in an online format.

5.1 Additional tools and functions to create more advanced tasks

5.1.1 Analysing concordances

The “Concordance” function is useful for finding language patterns for task design or for demonstrating certain language structures. More advanced language learners or corpus-experienced students could use this function to explore sentence structures on their own.

Select the “Word sketch” option and search for a word as explained in section 3.1.2.2. Open the concordance lines (cf. section 3.1.2.3). Within the concordance section, click on the “Collocations” symbol, i.e., the three dots at the top of the page (Fig. 20).

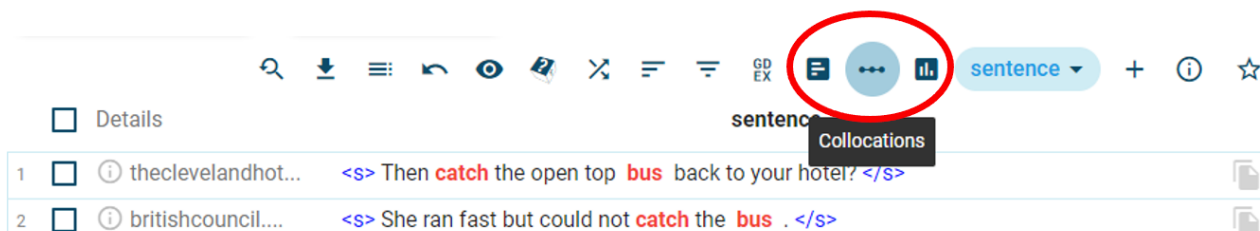


Fig. 20: The “Collocations” function

If you want to find the most frequent word forms right next to the keywords *catch + bus*, for instance, select the first position on the right side. Click “GO” (Fig. 21).

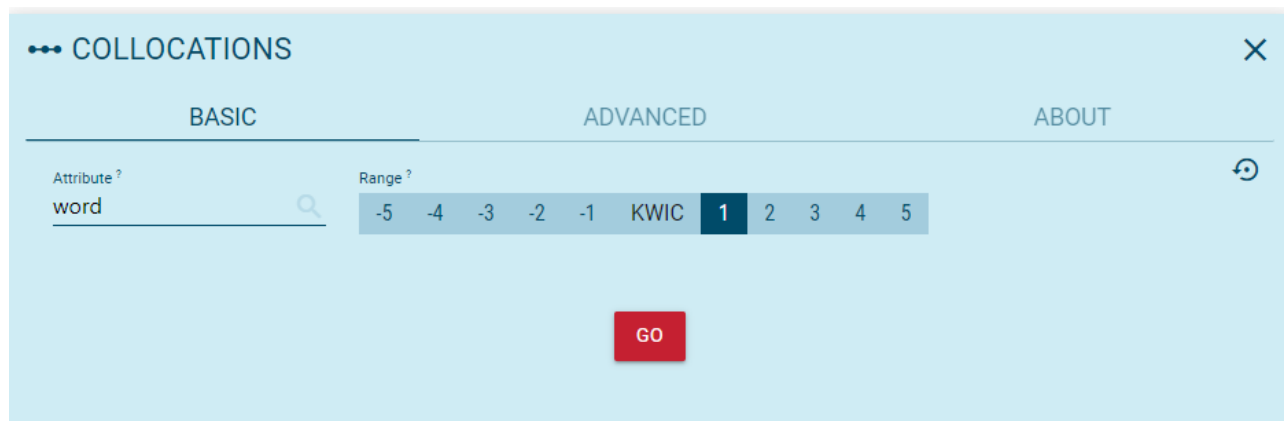


Fig. 21: Specifying the position of the collocates

You will then be presented with a list of the most frequent words directly following the collocation *catch + bus*. In Fig. 22, you can see that these are almost exclusively prepositions.

	Word	Cooccurrences [?]	Candidates [?]	T-score	MI	LogDice ↓	
1	<input type="checkbox"/> back	8	14,689,125	2.81	7.14	-5.81	...
2	<input type="checkbox"/> into	4	30,396,678	1.94	5.09	-7.86	...
3	<input type="checkbox"/> from	7	96,463,679	2.50	4.23	-8.72	...
4	<input type="checkbox"/> to	32	583,003,761	5.26	3.83	-9.12	...

Fig. 22: List of the most frequent words to the right

Once again, you can open the respective concordance lines by selecting one (or several) collocates and clicking on “show concordances” at the bottom of the page (Fig. 23)

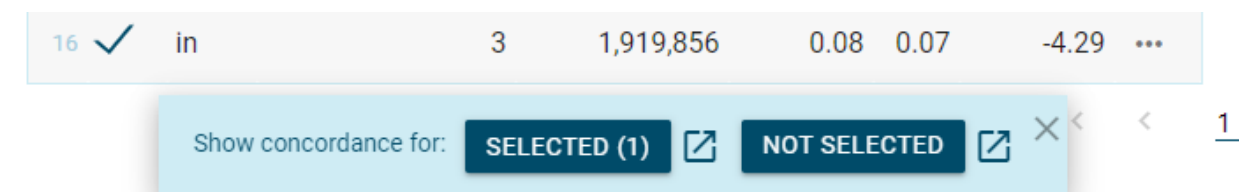


Fig. 23: Show concordances

5.1.2 Multi-word search

Go to the “Word sketch” function and search for a word (cf. section 3.1.2.2). Open the concordance lines (cf. section 3.1.2.3). Within the concordance section, click on the “Change criteria” symbol (Fig. 24).

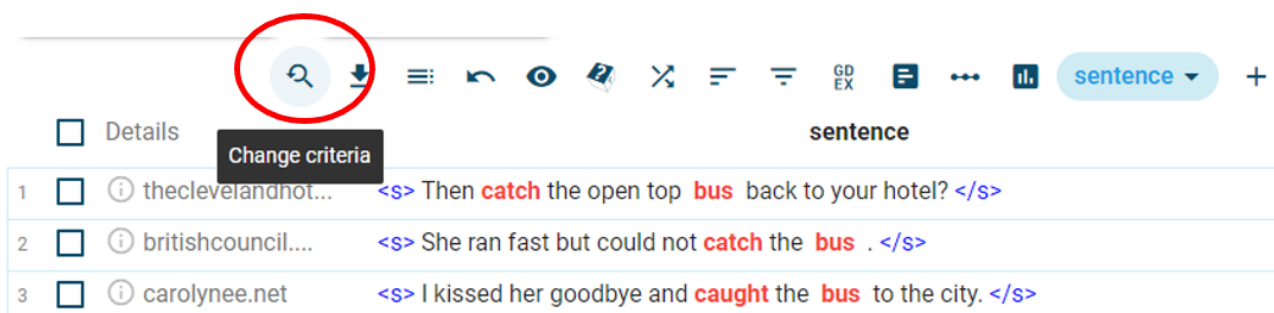


Fig. 24: Changing the search criteria

In the “Advanced search” settings, click on “Filter context” and tick “Part-of-speech context” (Fig. 25). This function allows you to search for an additional wordform that occurs in the same sentence as, e.g., *bus* and *take*.

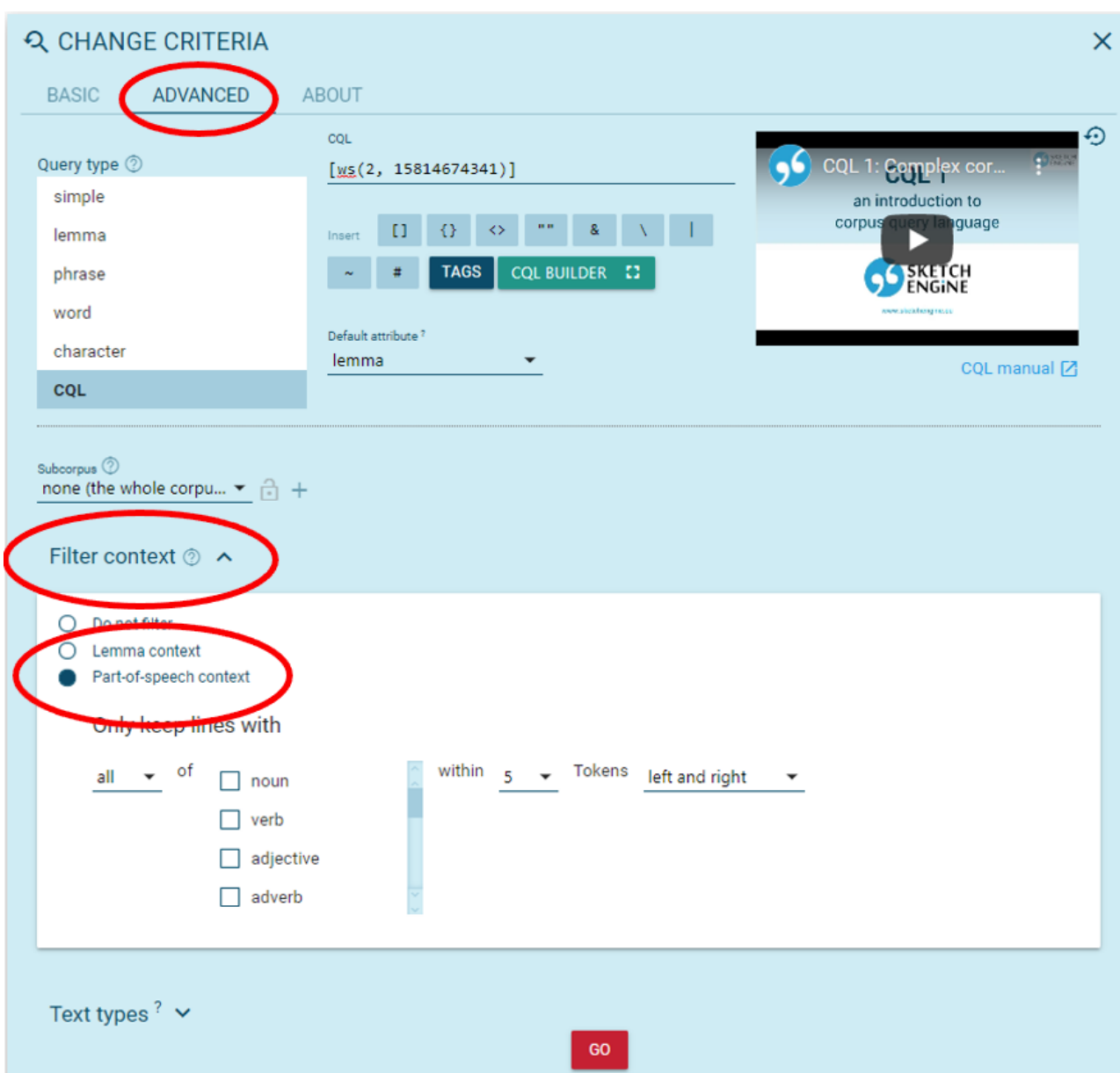


Fig. 25: Multi-word search

A drop-down menu for choosing the part(s) of speech you want to search for will appear. As an example, search for all prepositional phrases associated with the collocation *take + bus*. These are good exemplary phrases for further analysis. To do this, click “all” and tick “preposition” in the drop-down menu. To specify the position of the prepositional phrase relative to the keywords, select “1” Tokens “right” and click “GO” (Fig. 26). Sketch Engine will then only display concordances with prepositional phrases occurring immediately to the right of the collocation (Fig. 27).

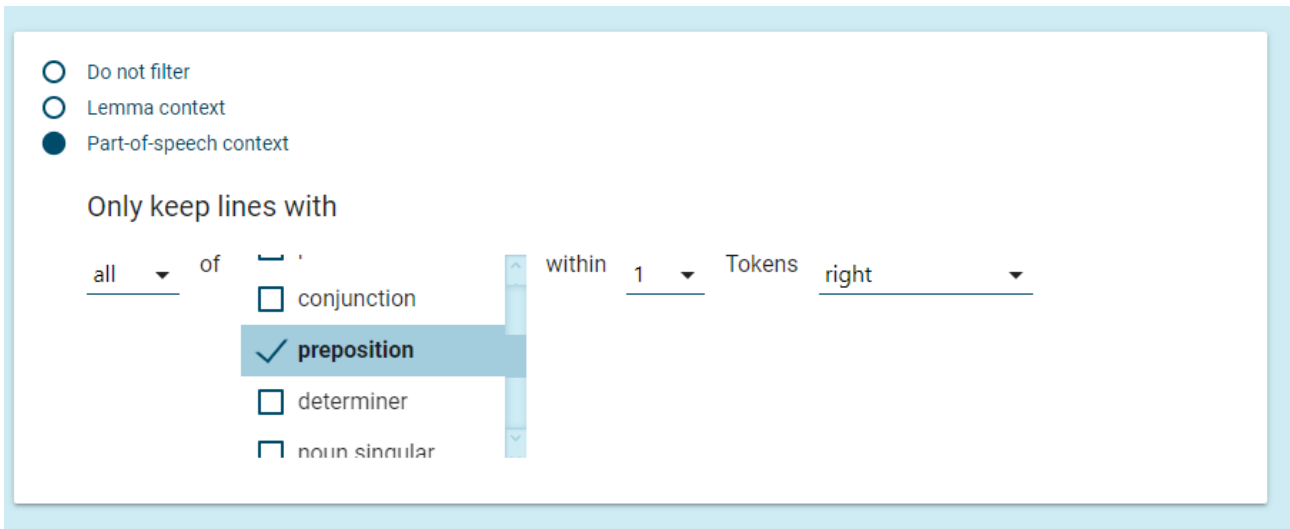


Fig. 26: Multiword search settings

	Details	sentence
1	<input type="checkbox"/> bluedogairporth...	<S> Catch the transfer bus to the terminal. </S>
2	<input type="checkbox"/> flash-game.net	<S> You should catch a bus as soon as possible. </S>
3	<input type="checkbox"/> athomeinthenort...	<S> The kids catch the bus in front of our house. </S>
4	<input type="checkbox"/> koreanwar-educa...	<S> I later went uptown to catch the bus for boot camp. </S>
5	<input type="checkbox"/> daltondatabank...	<S> From there I caught a bus to Aurora. </S>

Fig. 27: Multi-word search results

5.1.3 Visualisation

Another nice feature is the “Visualisation” function on Sketch Engine.

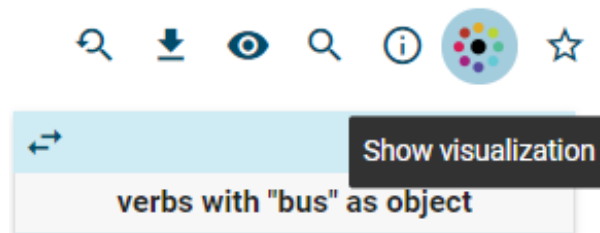


Fig. 28: Visualisation symbol

Visualising the word sketch (e.g., for demonstration purposes) can be easily done by way of the student-friendly, interactive visualisation tool on Sketch Engine. Open the word sketch interface and search for a word (cf. section 3.1.2.2). Click on the “Show visualization” symbol in the toolbar (Fig. 28). For a better overview, select only the categories you want to show and adjust the number of collocations you want to display until you are satisfied with the visualisation (cf. Fig. 29).

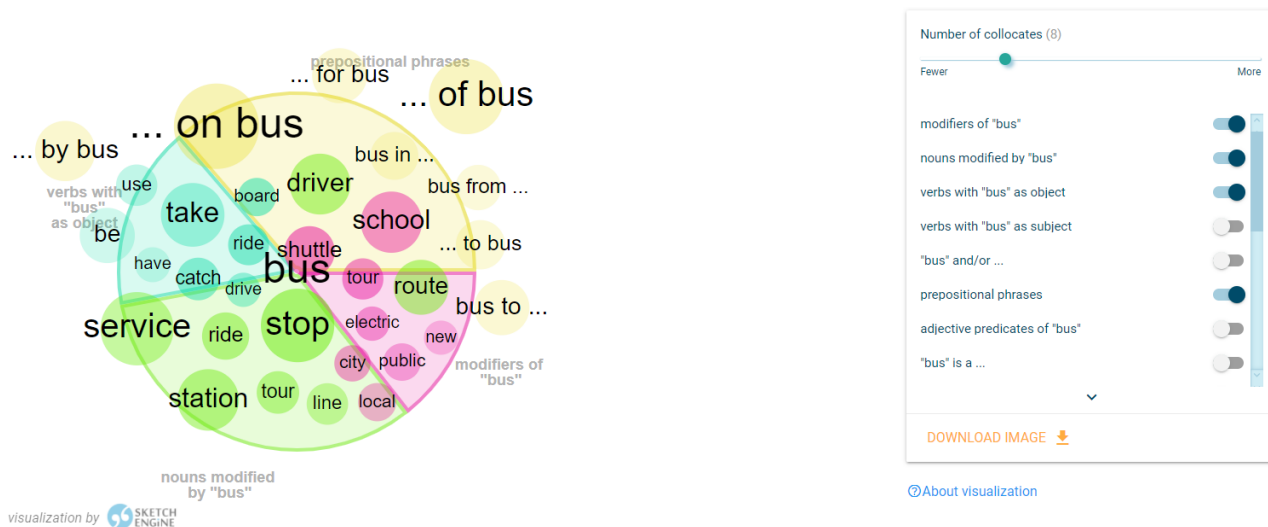


Fig. 29: Visualising the word sketch

The image is downloadable. Hover over one area with your cursor to highlight just one on the categories. This could help students formulate their own sentences, for example in task 3 of the lesson outlined in section 4.

6 Conclusion

Corpus-based task design is still primarily found in academic contexts. In this chapter, we intended to show how corpus tools can also be used by language learners of lower proficiency. The approach introduced in this chapter is applicable to many more topics and levels. The idea is to encourage learners to explore new word fields by creating word sketches and analysing words in context by themselves. Students actively engage with new vocabulary in context instead of simply memorising lists of isolated words from the textbook.

Word sketches like the ones created in this lesson are a good way of discovering authentic chunks of speech that a textbook could not provide. In this chapter, we have shown that SkeLL is a simple, yet practical tool that can give learners the means to becoming autonomous language learners, whilst Sketch Engine with its variety of useful functions has an enormous potential for teachers, especially when creating further exercises and visualisations.

7 Resources and references

Baer-Engel, Jennifer et al. 2017. *Green Line 4 G9*. Stuttgart: Ernst Klett Verlag.

Friginal, Eric. 2018. *Corpus linguistics for English teachers: New tools, online resources, and classroom activities*. New York, NY: Routledge.

Lancaster University. Corpus for schools: ELT Materials. <https://www.lancaster.ac.uk/corpusforschools/esl-teaching-materials/> (31 March, 2021).

Niedersächsisches Kultusministerium. 2015. Kerncurriculum für das Gymnasium, Schuljahrgänge 5-10: Englisch. <https://cuvo.nibis.de/cuvo.php?p=download&upload=139> (03 June, 2021)

SKELL – Sketch Engine for Language Learners. <https://SkeLL.sketchengine.eu/#home>.

Sketch Engine – Create and search a text corpus. www.sketchengine.eu.

3. "Let's have a party!": Teaching students how to express, accept and refuse invitations

SELKE TOMBRÄGEL

1 Introduction and rationale

Communicative approaches to teaching English as a foreign or second language aim to enable learners to master key communicative situations such as going to a restaurant, shopping, asking for directions, etc. As such, it is necessary to give students the opportunity to come into contact with authentic spoken language as opposed to the largely written-language-based and often unidiomatic materials presented in school textbooks. With this in mind, this chapter applies a communicative teaching approach by teaching students how to give, accept and refuse invitations in a socially accepted manner. This teaching objective is in line with our local secondary school curriculum, which states that students should be able to invite somebody and to respond to invitations (Niedersächsisches Kultusministerium 2015: 17). This chapter provides teachers with a detailed step-by-step guide on how to access and compile idiomatic spoken language material by conducting their own corpus linguistics research and how to process this material for teaching purposes. It thus shows how idiomatic spoken language use can form the basis of diverse communicative and meaningful learner activities, which prepare students for communicative encounters outside the classroom.

Outline and objectives

- Target learners: Lower secondary school (e.g. 8th grade)
- Teacher preparation time: ca. 60 min.
- Lesson time: ca. 90 min.

By the end of this lesson, students will be able to:

- Invite friends to a party
- Accept invitations in spoken English
- Politely refuse invitations in spoken English

2 Corpus, tools and methods

Quaglio (2009: 148-149) advocates the use of television dialogue as a surrogate for natural conversation, adding that "it offers a vast potential for pedagogical purposes". Accordingly, the Corpus of American Soap

Operas was selected for this lesson. It contains 100 million words of spoken American English from ten US soap operas from the early 2000s and is freely accessible via english-corpora.org. Note that, unlike many spoken corpora, which feature adult, uncensored spoken language, the Corpus of American Soap Operas is especially adequate for selecting language material for younger learners, as it does not include any swear words. To use the english-corpora.org, users will need to register. The free account only offers a limited number of searches per day but this is sufficient for teaching purposes.

The corpus research conducted by the teacher as part of the preparation for this lesson focuses on the KWIC functions of english-corpora.org. This function is particularly useful because it can help to counter the “over-emphasis in language teaching on single words” (O’Keeffe et al. 2007: 63); and to make students aware that “the meaning of a word is as much a matter of how it combines with other words in actual use” (Firth 1935 in O’Keeffe et al. 2007: 59).

3 Step-by-step guide

3.1 Teacher preparation – data collection

Step 1: Registering on english-corpora.org

To access the Corpus of American Soap Operas, connect to www.english-corpora.org and select the corpus by clicking on it (Fig. 1).

Corpus (online access)	Download	# words	Dialect	Time period	Genre(s)
iWeb: The Intelligent Web-based Corpus		14 billion	6 countries	2017	Web
News on the Web (NOW)		8.7 billion+	20 countries	2010-last month	Web: News
Global Web-Based English (GloWbE)		1.9 billion	20 countries	2012-13	Web (incl blogs)
Wikipedia Corpus		1.9 billion	(Various)	2014	Wikipedia
Corpus of Contemporary American English (COCA)		600 million	American	1990-2019	Balanced
Corpus of Historical American English (COHA)		400 million	American	1810-2009	Balanced
The TV Corpus		325 million	6 countries	1950-2018	TV shows
The Movie Corpus		200 million	6 countries	1930-2018	Movies
Corpus of American Soap Operas		100 million	American	2001-2012	TV shows

Fig. 1: Accessing the Corpus of American Soap Operas.

If you are curious to find out more about the different soaps included in the corpus, click on the “sheet-of-paper symbol” (Fig. 2).



Fig. 2: Further information on the corpus.

Step 2: Compiling corpus data with the “KWIC” function

The “Keyword in Context” (KWIC) function enables you to see in which context a keyword, occurs. First, type in the keyword, in this case *invite*. Second, select the words by which you want to sort your search. Select L for 1, 2, or 3 words to the left. Select R for 1, 2, or 3 words to the right. It is also possible, for example, to sort by one word to the left, then one and two words to the right. Note that you can click on the grey box with the asterisk to clear the entries and start over (Fig. 3).

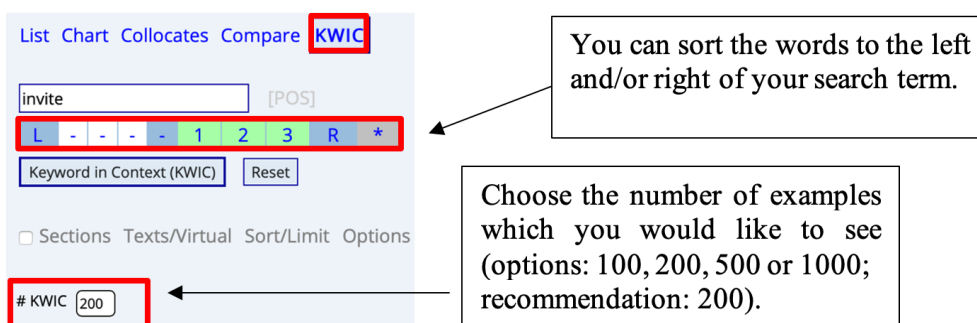


Fig. 3: Conducting a KWIC search.

As a result of this sample search, you will obtain a list of 200 examples of the verb *invite* and the three words to its right (Fig. 4). Elements which belong to the same structural category are highlighted in the same colour, which will help you and your students to identify common patterns of usage.

20	2009	YR	A	B	C	. Thanks . Victor: All right , why do n't you	invite	Eden to come out to the ranch one day ? Noah: Hey
21	2010	GH	A	B	C	told me you arranged this , asking her to call and	invite	Elizabeth out with the kids . Nikolas: I did . I did
22	2007	YR	A	B	C	if you do stay like this , I 'm gon na	invite	Elliot to come see you . Which may not be a bad
23	2006	DAYS	A	B	C	of stuff to do . We have to go around and	invite	everybody to the wedding tomorrow . Austin: You 're right .
24	2008	GL	A	B	C	me and then later we 'll throw a big party and	invite	everybody we know . But right now , this is just for
25	2006	GL	A	B	C	. # Marina: Yes , here is the catch : We	invite	everyone and anyone who knows and loves Harley . #
26	2010	ATWT	A	B	C	lot to do to get ready . Would you call and	invite	Faith ? Parker: Yeah , no problem . Liberty: Okay . #
27	2009	GH	A	B	C	for . Look at her . # Sam: You did not	invite	Franco into your life . You did n't ask him to fixate
28	2003	GL	A	B	C	know . I 'll have a family breakfast . I 'll	invite	Frank first . Last i heard , he was having a really
29	2006	ATWT	A	B	C	so cold . I 'll still have a few weeks to	invite	friends . I do have a big family . What 's wrong
30	2008	YR	A	B	C	? What did you think I was gon na do ,	invite	Harrison to Texas Hold ' Em ? Call for a bathroom break
31	2004	ATWT	A	B	C	And while we 're at it , why do n't we	invite	her detective sister , too ? Pilar: Why ? Simon: Oh ,
32	2010	GH	A	B	C	would give us a graceful out like call your sister and	invite	her for a visit . Yeah , well , is that what
33	2002	AMC	A	B	C	Did I ask who what ? Maddie: Mommy . Did you	invite	her for Christmas ? Edmund: Not yet . But the next time
34	2003	YR	A	B	C	right now . Brenda: You 're not that busy . Just	invite	her here for a drink at Marsino 's . A guy 's
35	2005	GH	A	B	C	, why do n't we just let her ? We 'll	invite	her in . Ric: Oh , that 's brilliant . Why did
36	2005	ATWT	A	B	C	got an even better idea . Why do n't we just	invite	her in ? # Emily: Where do you keep your blunt objects

Fig. 4: Results for the KWIC search of *invite*.

You may have observed that the search options, as displayed in Fig. 3, only returned the non-inflected verb form *invite*. In order to find representative examples including the different conjugated forms of the verb *invite* (e.g. he/she/it invites, he was invited) (Fig. 6), you need to search for the lemma *invite*. The lemma includes all conjugated verb forms such as *inviting*, *invited*, *invites* etc. To search for the lemma *invite*, type: *[invite]* or *INVITE* and click on “Keyword in Context (KWIC)” (Fig. 5).

List Chart Collocates Compare **KWIC**

[invite] [POS]

L - - - - 1 2 3 R *

Keyword in Context (KWIC) Reset

Fig. 5: Lemma search for the verb invite.

22	2006	ATWT	A	B	C	your mom would love that . Will: Well , I 'll invite her along for the ride . Gwen: If you do that ,
23	2006	BB	A	B	C	Donna: I thought they gave her over half the company ? Invited her back into their family with open arms ? # Nick: Just
24	2005	GH	A	B	C	into has nothing to do with me . I did not invite her down here to go snooping through my business . Skye: Oh
25	2003	ATWT	A	B	C	did I listen to you ? Why did I think that inviting her into this baby 's life would help anything ? If she
26	2007	YR	A	B	C	I do n't know , maybe . I was thinking about inviting her over here for a while , just to make sure that
27	2010	GH	A	B	C	And rather than just writing her a check , he 's invited her to come and stay and said bring your sister with you
28	2007	DAYS	A	B	C	. Nick: What ? No . Chelsea: Well , should I invite her to join us , or are you gon na do it
29	2010	ATWT	A	B	C	. Chris: Why do n't you just make a gesture ? Invite her to the anniversary party , all right ? # Henry: Can
30	2004	YR	A	B	C	my daughter correctly , you took advantage of the moment to invite her to your ranch ? # Victor: What do you mean ,
31	2002	YR	A	B	C	But , come on . Isabella: What 's the harm of inviting her ? She 's not going to come anyway . Paul: I
32	2005	ATWT	A	B	C	taking Mike so long with the right ones . Should we invite him for dinner ? # Henry: Yeah , sure , we can
33	2004	BB	A	B	C	know what I do n't understand ? Why you had to invite him here in the first place . # Nick: He 's trying
34	2008	BB	A	B	C	That was really creepy . And then he insisted that I invite him in . Ridge: What ? Ashley: He did n't force the
35	2011	BB	A	B	C	Okay , well , why do n't you call Liam and invite him over ? Maybe he 'd like some . Amber: Because he
36	2008	BB	A	B	C	Brooke: You know , I want to talk to Rick and invite him to our rehearsal dinner . Ridge: Not a good idea at
37	2002	GH	A	B	C	, everybody . Felicia: Oh , hello . Ned: Who thought inviting him was a good idea ? Edward: It 's always wise to
38	2006	BB	A	B	C	? Taylor: Hector . Stephanie: Oh . Taylor: I did n't invite him . He just showed up to wish me happy new year

Fig. 6: Results for lemma search invite.

Step 3: Searching for particular parts-of-speech

When looking for similar patterns, it is important to search for *invite* as either a noun or a verb in order to obtain comparable examples. This selection can be made by clicking on the light grey button "POS" (= Part-of-Speech) next to where the search term is entered (Fig. 7). A drop-down menu opens and "verb.all" or "noun.all" can be selected.

List Chart Collocates Compare **KWIC**

invite_v* **verb.ALL**

L - - - - 1 2 3 R *

Keyword in Context (KWIC) Reset

Sections Texts/Virtual Sort/Limit Options

KWIC 200

Fig. 7: Searching for a specific part of speech.

In the following you will find some results for exemplary searches of *invite* as a verb.

27	2011	YR	A	B	C	crowd her . Bro . Neil: No . I should just	invite	her	to	a	wedding	where her husband was killed . Malcolm: Okay
28	2008	YR	A	B	C	to find out if she 's really sincere . Let 's	invite	her	to	dinner	, I mean , for real . Will you	
29	2010	ATWT	A	B	C	remember not too long ago that you two would n't even	invite	her	to	Hallie 's	christening . Gwen: Okay , your mom is	
30	2005	ATWT	A	B	C	doing today . # Jack: Well , why do n't you	invite	her	to	join	us ? Ask Emma , I 'm sure she	
31	2011	AMC	A	B	C	hire somebody full-time to take care of our baby , you	invite	her	to	move	into	our house without discussing it with me .
32	2009	GH	A	B	C	either refrain from inviting her to Wyndemere all together or	invite	her	to	spend	the night . Having her for dinner only does	
33	2002	DAYS	A	B	C	be further from the truth . John: Then why did you	invite	her	to	the	dinner	party ? Tony: I thought I explained it
34	2009	GL	A	B	C	it 's ringing ? What ... you called Natalia ? Emma:	invite	her	to	the	party	. Please ? Olivia: I 'll mention it

Fig. 8: Pattern *invite sb to sth*.

45	2001	PC	A	B	C	some guy . He was lost . Kevin: You did n't	invite	him	in	for	green	tea and offer to take him where he
46	2001	PC	A	B	C	here earlier . Michael: What ? You -- you did n't	invite	him	in]	did	you ? Eve: No . Of course not
47	2001	PC	A	B	C	Jack: Because Caleb ca n't come in if you do n't	invite	him	in]	All	right ? That 's one of the rules

Fig. 9: Pattern *invite sb in*.

156	2008	ATWT	A	B	C	family room , too . Lily: Carly , I did n't	invite	you	over	here	to	trade decorating tips . Carly: I know .
157	2007	AMC	A	B	C	Thank you . You know what ? I did n't just	invite	you	over	here	to	talk about it , because I want you
158	2001	OLTL	A	B	C	. # Antonio: So I thought I 'd call you and	invite	you	over	to	my place for lunch and , you know ,	
159	2009	OLTL	A	B	C	see you guys out somewhere , or whether my mom would	invite	you	over]	and we would finally meet . But I decided	

Fig. 10: Pattern *invite + sb + over*.

Fig. 11: Search settings for searching for *invite* as a noun.

Now also search for the nouns *invite* and *invitation*.

76	2008	AMC	A	B	C	through . Yeah , that 's right . You have the	invite	list	?	Good	make	sure Zach Slater is coming , too
77	2010	AMC	A	B	C	Greenlee: Why mess with perfection ? Oh , and about the	invite	list	?	want	one	sent to Kendall and Zach . I
78	2008	YR	A	B	C	is n't your party , is it ? Colleen: With this	invite	list	?	No	Not	a chance . Have fun , though

Fig. 12: Pattern *invite + list*.

List Chart Collocates Compare **KWIC**

invitation_nn* noun.ALL

L - - - 1 2 3 R *

Keyword in Context (KWIC) Reset

Sections Texts/Virtual Sort/Limit Options

Fig. 13: Search settings for searching for invitation as a noun.

9	2004	ATWT	A	B	C	And I'm pretty sure I'm gon na get and invitation for breakfast from Dad and Kim . Margo: Yeah , I'm
10	2008	ATWT	A	B	C	Holden: I think polo is abusive . It 's an open invitation for horses to get injured . Neal: Well , the same could
11	2002	ATWT	A	B	C	? Craig: You think it 'd be possible to wrangle an invitation for Lucy 's old man ? Huh ? Huh ? Holden: This
12	2003	ATWT	A	B	C	knew that it was from Travers . It was his little invitation for me to go see him . Jessica: Why did n't you
13	2007	GL	A	B	C	any longer . I can not believe you turned down my invitation for some watery fun . Hey , you still thinking about Griggs
14	2002	PC	A	B	C	find , no buts about it . I have a personal invitation for this evening ; so with or without the two of you
15	2007	BB	A	B	C	to keep our stockholders happy . Jackie: You received the invitation for tomorrow 's show ? George: Yes , ma'am . Vijay and
16	2010	BB	A	B	C	No , I ... (Laughs) I was n't fishing for an invitation for tonight ; I ... (chuckles) Justin: Donna , I would be
17	2010	OLTL	A	B	C	Nora when Matthew sued them . So I 've had my invitation for weeks ; # Dani: A little to the left . Yeah
18	2009	YR	A	B	C	I-I took it when I saw this was an -- an invitation for you to go up to the cabin . She was trying
19	2004	AMC	A	B	C	faked . Look at her . Greenlee is just an open invitation for you to have your way with her . So if at
20	2003	DAYS	A	B	C	Hey , guys . # Belle: Rex ; I have an invitation for you ; Whoops . Ha ha ha ha . # Rex:

Fig. 14: Search results for searching for the noun invitation.

Step 4: Documenting research results

Having familiarized yourself with the KWIC feature of the corpus and skimmed through the examples, you now need to select data suitable for your lesson which exemplify the speech act of "inviting somebody". English-corpora.org offers you the useful option of creating your own lists containing only the examples which are most interesting or relevant for your teaching purposes. To save individual examples, simply click on all the examples that you would like to use in class (they will be highlighted in blue), give your list a name in the textbox (Number 4 in Fig. 15) and click on "SAVE LIST". You can also add to lists that you have previously created by selecting the list name in the dropdown menu (Number 3 in Fig. 15).

It is possible to save KWIC entries and to access and modify these lists at a later time.

1 CLICK FOR MORE CONTEXT			2 <input type="checkbox"/> [?]			SAVE LIST	5 CHOOSE LIST	3 CREATE NEW LIST	<input type="text"/>	4
1	2011	MAG	Source Info 1	7	A	B	C	the entry will be displayed here; click on it to add it to the list 6			
2	2003	ACAD	Source Info 2		A	B	C	the entry will be displayed here; click on it to add it to the list			
3	2001	FIC	Source Info 3		A	B	C	the entry will be displayed here; click on it to add it to the list			
4	1998	NEWS	Source Info 4		A	B	C	the entry will be displayed here; click on it to add it to the list			

1. Click to see an expanded display (about 100 words), as well as full source data for the entry
2. Select all of the entries in the list.
3. If you have already created a saved list and want to add to that list, select it.
4. If you want to create a new saved list, enter the name.
5. See the Saved List. At that page you can also delete entries, move entries between lists, and expand many entries at one time.

To select entries:

6. Click on one of the entries to select or de-select it.
7. You can also categorize the entries within a list by choosing A, B, or C. For example, you might want to sort the entries into one of three different groups, depending on their meaning or usage.

Fig. 15: Saving lists of concordance lines.

One option is to save lists of useful frequent structures, e.g. *invite + me/him/her/them + over/to*, so that your students may inductively infer the underlying patterns based on the similarity of the examples. You can access your lists anytime. Choose the list you would like to access and click on "SAVE LIST" (Fig. 16).

(SHUFFLE)

CLICK FOR MORE CONTEXT			[?]			SAVE LIST	CHOOSE LIST	Example	CREATE NEW LIST	<input type="text"/>	[?]
								invitealong				
1	2012	YR	A B C	Well put. Harmony: (Sighs) Anyway, listen, I-it was kind of you to invite me along. I-I appreciate it, but I plan to stay far away from								
2	2009	GH	A B C	You would have a blast. Which is -- that's probably my cue to invite you along. Johnny: No -- no, I get it. I understand why								
3	2008	ATWT	A B C	have dinner? Henry: All right. Mwah. Oh, hey, we should invite Maddie along to dinner. We haven't had a chance to celebrate her being								
4	2008	GL	A B C	want to figure out what they should do on their own. But they will invite us to go along for the ride, right? I mean, because we								
5	2007	YR	A B C	(Laughs) Gloria: Kevin, please don't be rude. Jeffrey was kind enough to invite you along. Don't give him the third degree. Kevin: Well, I								

Fig. 16: Accessing a saved list.

[?]			#	<input type="text"/>	-	<input type="text"/>	[?]	● DELETE ENTRIES ● MOVE ENTRIES ● EXPAND ENTRIES [?]				
1	<input type="checkbox"/>	SOAP:2009 GH	You would have a blast. Which is -- that's probably my cue to invite you along. Johnny: No -- no, I get it. I understand why									
2	<input type="checkbox"/>	SOAP:2008 GL	want to figure out what they should do on their own. But they will invite us to go along for the ride, right? I mean, because we									
3	<input type="checkbox"/>	SOAP:2006 ATWT	Oh, yeah, your mom would love that. Will: Well, I'll invite her along for the ride. Gwen: If you do that, I will divorce									

Fig. 17: The "saved list" for invite + along.

Alternatively, you may copy the examples and/or the expanded contexts into an empty Word document to use this data for the preparation of teaching material. Either way, you should document your research results as it forms the basis for preparing your teaching materials.

Step 5: Categorizing research results

Based on your research results, you can now categorize the examples you selected according to your teaching objectives. The table below presents such an example categorization.

Table 1: Corpus-derived patterns of usage for *invite* and *invitation* sorted by learning goals.

Give/express an invitation	Accept an invitation	Refuse an invitation
<ul style="list-style-type: none"> • I would like to invite <i>sb.</i> • Could + invite • Invite <i>sb.</i> back • Invite <i>sb.</i> in • Invite <i>sb.</i> to <i>sth.</i> • An invitation to <i>sth.</i> • Invite <i>sb.</i> to <i>sth.</i> • Invite <i>sb.</i> for <i>sth.</i> • Invite <i>sb.</i> to join... • To invite <i>sb.</i> over 	<ul style="list-style-type: none"> • To thank <i>sb.</i> for the invite • I appreciate the invite • Thank you, I accept the invite • Thanks for inviting me/us • To accept an invitation 	<ul style="list-style-type: none"> • Thank you for the invite, but I cannot... • To turn down an invitation • I appreciate the invite, but... • Thank for the invite, but I'm going to pass...

3.2 Teacher preparation – designing corpus-based teaching materials

Warm-up

Step 6: Selecting a video scene of an invitational act

For the **warm-up activity**, select an appropriate invitation scene, ideally from one of the soaps in the corpus. The expanded context information of a specific concordance line as illustrated below (Fig. 18), includes information about the date on which this specific scene was first broadcasted. Therefore, you may try to search the internet to find this specific episode, though this can be a time-consuming task! Alternatively, search for an alternative invitation scene on YouTube, e.g. from the soap “All my children” (on YouTube, search for: “All my children invitation”): <https://www.youtube.com/watch?v=3awzuCGBqqo>.

Source information:

Date	2009 (2009-07-02)
Title	General Hospital
Source	General Hospital

Expanded context:

family are really something. It's, like, classic old-school Italian, you know? The big meal in the middle of the day, and when I say big, I mean, like, roman emperor's feast. It's a wonder we're not all 400 pounds. Cousin Violetta's pushing it. Johnny: Are they all as outgoing as you are? Olivia: Outgoing? That's a nice euphemism. Yeah, it's pretty much wall-to-wall type A personalities. You would have a blast. Which is -- that's probably my cue to **invite you along**. Johnny: No -- no, I get it. I understand why you can't. # Alexis: We need to discuss what happened here. Kristina: Sure, Alexis: Would you mind turning that off and looking at me, please? I understand how hurt you are, and how disappointed you are in me. You have every reason to be. But you shouldn't have shown Molly that blog, because she's too young to understand the hateful things they said about me. You're

Fig. 18: Broadcasting information.

Pre-task

Step 7: Preparing playing cards

This interactive **pre-task** is designed to focus the students' attention onto the diverse contexts, in which the two most basic terms linked to the topic of invitation (*invite* and *invitation*) may occur. To create the pre-task material, access your list(s) of examples which you have saved on the English-corpora website (step 4) or as the Word document you created. Split a selection of these sentences into meaningful chunks of two to five words. Print or write the individual chunks onto large, colourful cards.

Table 2: Example playing cards to be pinned onto the board.

I would very much like	to invite you	to come
Thanks	for the invite	but I can't stay

First task

Step 8: Designing a worksheet based on the KWIC material

For the **first task**, you need to design a worksheet based on the corpus data which you compiled through your KWIC research (Step 2). Basically, you provide the students with blocks of examples which illustrate recurrent patterns.

**Worksheet „Working with Corpus Data“
Task 1**

a) Below you find some examples which were taken from US American Soap Operas (spoken language). Look at each block of examples and try to identify a pattern, which the examples have in common. Write down your ideas.

Block A:

27	2011	YR	A	B	C	crowd her . Bro . Neil: No . I should just invite her to a wedding where her husband was killed . Malcolm: Okay
28	2008	YR	A	B	C	to find out if she's really broken . Let's invite her to dinner . I mean , for real . Will you
29	2010	ATWT	A	B	C	remember not too long ago that you two would n't even invite her to Halle's christening . Gwen: Okay , your mom is
30	2005	ATWT	A	B	C	doing today . # Jack: Well , why do n't you invite her to join us ? Ask Emma , I'm sure she
31	2011	AMC	A	B	C	hire somebody full-time to take care of our baby , you invite her to move into our house without discussing it with me .
32	2009	GH	A	B	C	either refrain from inviting her to Wyndemere all together or invite her to spend the night . Having her for dinner only does
33	2002	DAYS	A	B	C	be further from the truth . John: Then why did you invite her to the dinner party ? Tony: I thought I explained it
34	2009	GL	A	B	C	it's ringing ? What ... you called Natalia ? Emma: Invite her to the party . Please ? Olivia: I'll mention it

Pattern: _____ + _____ + _____ + _____

Block B:

156	2008	ATWT	A	B	C	family room , too . Lily: Carly , I did n't invite you over here to trade decorating tips . Carly: I know .
157	2007	AMC	A	B	C	Thank you . You know what ? I did n't just invite you over here to talk about it , because I want you
158	2001	DLTL	A	B	C	. # Antonio: So I thought I'd call you and invite you over to my place for lunch and , you know ,
159	2009	DLTL	A	B	C	see you guys out somewhere , or whether my mom would invite you over and we would finally meet . But I decided

Pattern: _____ + _____ + _____

Block C:

37	2009	ATWT	A	B	C	, that's up to her . Roxanna: How can I turn down an invitation from my favorite boy ? Craig: All right .
38	2008	GH	A	B	C	ago . # Nikolas: You made it . Nadine: I never turn down an invitation to ride . Nikolas: Oh , I'm glad
39	2006	AMC	A	B	C	. # Julia: Note to self -- when someone tries to turn down an invitation I hear them out . Ryan: It was incredibly

Pattern: _____ + _____ + _____ + _____

Further examples should be added

b) After having discovered how to use the nouns “invitation” and “invite” as well as the verb “invite”, try to relate the useful phrases with the correct category.

Thank sb for + invite / should have a party / Thank you, I accept the invite / Thank you for the invite, but I'm gonna pass... / invite sb in / invite + back / could + invite / turn down an invitation / I appreciate the invite / I appreciate the invite, but... / thanks for inviting sb / invite sb to sth / invite sb for sth / invite sb to join / accept an invitation / invite to sth / invite + noun / I would like to invite sb / thank you for the invite, but I can't... / to invite sb over / an invite to sth

Accept invitation	Give/express an invitation	Refuse an invitation

Fig. 19: Example worksheet for the first task.

Please note that Fig. 19 is a sample worksheet and that you should add further examples.

Second task

Step 9: Designing role-play cards

While the first task focused the students' attention on patterns of usage, the **second task** aims to motivate students to actively use these patterns in communication. In order to achieve this aim, it is recommended to create role-play cards. The corpus data you collected (in the form of concordance lines and their expanded contexts) offer numerous example situations which you can use for creating the cards. You can either directly copy the "expanded context" onto the cards if the situation is clear enough that way, or rephrase and summarize the depicted scene in a couple of sentences if necessary. While the situation is presented on the back of the card, the front side states the topic of the situation e.g. "Refusing an invitation" (Table 3).

Table 3: Example role-play cards.

<p>– front –</p> <p>Inviting somebody</p>	<p>– back –</p> <p>Situation:</p> <p>You have decided to celebrate your birthday, which is coming up the following week. On your way to soccer practice you meet a good friend and immediately want to invite him/her.</p>
<p>– front –</p> <p>Refusing an invitation</p>	<p>– back –</p> <p>Situation:</p> <p>Your best friend invites you to his/her birthday. Unfortunately, you will be on vacation with your parents.</p>

Third task

Step 10: Formulating the final task of the session

The final task of the session requires the students to practise using the structural patterns and phrases they have learned over the course of the session in writing by designing an invitation card (Fig. 20).



Fig. 20: Example Invitation Card Task (<http://pixabay.com>).

Task 3

Write an invitation card for an imaginary party. Don't forget to mention:

- What you are celebrating
- The date, time and place
- What the guests should bring
- Information on how to respond to your invitation.

Lesson plan

Table 4: Lesson plan 'Let's have a party!'

Phase	Task	Implementation	Interaction form	Material/Medium
Warm-up 10 min	Video impulse "invitation"	Quick impulse Video Discussion on the video	plenum	YouTube video
Pre-task 20 min	Guess what comes before and after!	Reorganizing cards on the board Discussion of results	group work/ plenum	Cards for the board (and/ or groupwork)
Task 1 20 min	Hands-on KWIC	Working with concordance lines, discovering structural patterns and useful phrases Comparing results	Individually/ partner work/ plenum	worksheet
Task 2 25 min	Role-Play	Practicing new structures and phrases in communication Presentation of mini role plays	partner work plenum	Role-play cards
Task 3 15 min	Designing an invitation card	Writing an invitation card! Homework: finish it!	individually	Sample card via document camera

Instructions

Pre-Task: Language chunk cards

The idea is to pin a number of sentence "chunks" onto the board. In a class of 24 students, form four groups of six students. Within their groups, the students discuss and guess which of the chunks presented on the board belong together. Each group then sends one group member to the board to reorganize the chunks in the correct order. The group results are then discussed in class and secured on the board.

First task: Discovery learning of key patterns

In this task, students work through the worksheet designed in 3.2.3 individually or in pairs. Depending on the class, time resources and classroom facilities, this task could also be redesigned as a hands-on Data-Driven-Learning (DDL) task. In this case, the students themselves work online with the corpus. This, of course, means that they need access to computers and should have received a general introduction to corpus research similar to the one provided for teachers in this chapter. If these requirements can be met, the students work "hands-on" with the corpus in order to discover structural patterns. Instead of being given the pre-selected corpus data, they can investigate and select examples sentences themselves.

Second task: Role-plays

The class is split into pairs of two students: each pair receives approximately ten role-play cards; the students then have to act out the described situation by choosing appropriate phrases/patterns to talk about invitations. They can use the worksheet; especially task 1b (see worksheet Fig. 20) may help them to decide which structure to use in which situation.

Third task: Invitation cards

This task can conclude the lesson or, alternatively, be set as homework. Weaker students may be provided with an example invitation text to support them in this task.

4 Options and further ideas

As an extension to the invitation card writing task, students can “invite” each other to their parties and then accept or refuse invitations on the basis of a roll of the dice. The number on the dice can be suggested by the students and then summarized on the board.

Possible key to the die game

- 1 = Accept the invitation.
- 2 = Accept the invitation and ask if your friend/sibling can also come.
- 3 = Accept the invitation and offer to bring something.
- 4 = Accept the invitation but warn that you can only come an hour later.
- 5 = Refuse the invitation. You have something else planned.
- 6 = Refuse the invitation. You are already invited to another party.

Further follow-up lessons could focus students' attention on the degree of politeness in various invitation scenes, particularly with regard to the modal verbs used to express politeness. The following combinations are particularly frequent in the corpus:

- *Would + invitation*
- *Would love to + invite*
- *Couldn't + invite*
- *May + invite*
- *Should + invite*
- *Should have a party*
- *Ought to + invite*

The following search (Fig. 21) can help to find the most frequent modal verb + *invite* patterns.

List Chart **Collocates** Compare KWIC

_vm Word/phrase verb.MODAL

invite Collocates [POS]

+	4	3	2	1	0	0	1	2	3	4	+
---	---	---	---	---	---	---	---	---	---	---	---

Sections Texts/Virtual Sort/Limit Options

Fig. 21: Finding modal verbs in the context of invitations.

Task "Modal verbs"

In the corpus examples below, the speakers often use specific words to express what they are saying in a specific manner. Can you find these words?

Knock on door. Alexis: *Hi*. James: *Hello*. *I believe these belong to you*. Alexis: *Oh, thank you. Thank you very much, I really appreciate it. I would invite you in. I'm just not feeling that well right now.* (General hospital; 2007-04-09)

Jonathan: *Would you call her, please? Please tell her she doesn't have anything to worry about. I — I want her to concentrate on having a good time tonight.* All my Children(2007-05-02)

Luke: *Mmm. You know what? If I do say so myself, these pancakes were awesome.* Faith: *Why couldn't we invite Dad?* Lily: *Dad had a lot to do at the farm.* Faith: *Mom, why won't you tell us what's going on? Is Dad mad at you again?* (As the World Turns; 2008-07-31)

Hope: *I think that you should invite her back for dinner one night.* Liam: *I think she will not like that idea at all.* Hope: *I don't think she will, either.* (Bold and Beautiful 2012-04-24)

Katie: *Right. Right, we should have a party.* Jack: *Exactly, I was thinking more of a cookout. Here. Later today.* Katie: *Today?* Jack: *Today.* Katie: *Well, it's kind of last minute notice...* (As the World Turns 2007-08-30)

Casey: *You want to go grab lunch?* Alison: *Actually, yes, I would. I worked the early shift, so I'm done for the day.* Casey: *Excellent. I'll take you out. My treat.* Alison: *Well, how can I refuse an offer like that? Should we invite Jade?* As the World Turns (2009-04-30)

Carly: *Well, I guess we'll have a little more bubbly while we wait. What shall we toast to this time?*
 Craig: *Uh, how about to the most beautiful, compassionate woman I know?* Carly: *You ought to invite her to dinner.*
 (As the World Turns 2009-05-14)

How to express:

5 Caveats and limitations

Difficulties might occur when working with the Corpus of American Soap Operas via the interface english-corpora.org for the first time. Non-paying users will regularly receive a notification to upgrade their account (Fig. 22).

简体中文 中文 한국어 Русский العربية	<p>There is no cost for basic access to English-Corpora.org. But you will occasionally see this message, which asks you to consider upgrading to a "premium" account.</p> <p>if you have a premium account (\$10.00 per month, \$30.00 for one year, \$55.00 for two years, or \$75.00 for three years), then you will not see this message anymore. You will also have increased access to the corpora, more features, and you will help to support English-Corpora.org.</p> <p>if you are a student, your school or university can also get an academic license, and all of you will have the benefits of a premium account.</p> <p>Thanks for your support.</p>
---	---

Fig. 22: Notification to upgrade.

This notification may be disregarded, but requires a little patience. Having waited for a few seconds, users then have the option to continue their search without upgrading (Fig. 23).

[Click here to continue with your search](#)

简体中文
中文
한국어
Русский
العربية

There is no cost for basic access to English-Corpora.org. But you will occasionally see this message, which asks you to consider upgrading to a "premium" account.

if you have a **premium account** (\$10.00 per month, \$30.00 for one year, \$55.00 for two years, or \$75.00 for three years), then you will not see this message anymore. You will also have increased access to the corpora, more features, and you will help to support English-Corpora.org.

if you are a student, your school or university can also get an **academic license**, and all of you will have the benefits of a premium account.

Thanks for your support.

Fig. 23: How to continue a search without agreeing to upgrade.

6 Conclusion

Even though the initial preparation of this lesson based on self-compiled corpus data is more time-consuming than ordinary lesson preparation, this chapter has shown that the time investment is worth the effort. The preparation time will continuously decrease as teachers become more experienced in working with corpora. Once collected, the corpus data can be reused numerous times and the procedure can easily be applied to any other topic which aims at teaching students how to master communicative situations; it is even applicable to other foreign languages. Teachers of Spanish, for example, may want to consult www.corpusdelespanol.org, whose interface will look very familiar to experienced users of english-corpora.org since it is based on the same software.

7 Resources and references

AMCthroughtheyears. 2014. All My Children – 1999 – An Invitation to the Crystal Ball. *YouTube*. <https://www.youtube.com/watch?v=3awzuCGbqqo> (3 December, 2020).

Baer-Engel, Jennifer, Carolyn Jones, Harald Weisshaar, Alison Wooder, Cornelia Kaminski, Elise Köhler-Davidson, Katja Krey, Friderike Ulmer & Bernd Wick. 2017. *Green Line 4 [Schülerband]: Für Klasse 9 an Gymnasien (Bundesausgabe Ab 2014)*. Harald Weisshaar (ed.). Stuttgart; Leipzig: Ernst Klett Verlag.

Corpus of American Soap Operas. <https://www.english-corpora.org/soap/>

Harger, Laurence & Cecile Niemitz-Rossant. 2016. *English G – Access. 4, [Schülerband]*. Jörg Rademacher (ed.). Berlin: Cornelsen.

Neveling, Christinae. 2016. Verfügung über sprachliche Mittel. In Bär, Marcus, and Manuela Franke (eds.), *Spanischdidaktik: Praxishandbuch für die Sekundarstufe I und II*, 61-91, 237-249. Berlin: Cornelsen.

Niedersächsisches Kultusministerium. 2015. Kerncurriculum für das Gymnasium, Schuljahrgänge 5-10: Englisch. <https://cuvo.nibis.de/cuvo.php?p=download&upload=139> (09 January, 2020).

O'Keefe, Anne, Michael McCarthy & Ronald Carter. 2007. *From corpus to classroom: Language use and language teaching*. Cambridge; New York: Cambridge University Press.

Oxford University Press. 2003. *Das große Oxford Wörterbuch*. Oxford: Oxford University Press; Cornelsen.

Quaglio, Paulo. 2009. *Television dialogue: The sitcom Friends vs. natural conversation*. Amsterdam; Philadelphia: John Benjamins.

PART III

LESSON IDEAS: UPPER SECONDARY SCHOOL

Corpus-informed EFL lessons for upper secondary school

- Differences in the grammatical structures of 'tell' and 'say' by **Geeske Dünnhaupt** and **Lena Haase**
- How to do data-driven learning (DDL) with secondary school pupils to improve their writing by **Katharina Busmann**
- Using corpora in the EFL classroom for creative writing by **Jonas Dammann**
- Teaching neologisms using corpus tools by **Sarah Winkelmann** and **Nicol Sperling**
- Commas in argumentative writing by **Marie Nottmeier**, **Alina Sophie Peters** and **Lara Warnecke**

4. Differences in the grammatical structures of 'tell' and 'say'

A data-driven learning approach

GEESKE DÜNNHAUPT AND LENA HAASE

1 Introduction

The verbs *say* and *tell* have a similar meaning and, in German, both can be translated as *sagen*. As a result, German-speaking learners of English tend to use the grammar of these words synonymously and frequently produce unidiomatic phrases like *she told to me* or *my teacher said me that*. In fact, this problem is not confined to German-speaking learners of English: it is common to many Learner Englishes – as attested, for example, in the *Longman Dictionary of Errors* (Turton & Heaton 1997). To counter this difficulty, the following lesson focusses on the different grammatical structures of the two verbs, which the students will ascertain from authentic language samples. Students benefit from a corpus-based approach to learning about these two reporting verbs because corpora provide them with authentic language material. Knowing how to work with corpus analysis tools and being able to interpret the extracted data can positively contribute to students' learning strategies and learner autonomy.

Justification for applying such an approach can be found in the curriculum of Lower Saxony as it encourages teachers to include a variety of learning strategies in their teaching:

Lernstrategien beziehen sich auf den Erwerb der sprachlichen Mittel und bezeichnen die Strategien, die ein Lerner zunehmend selbstständig und individuell unterschiedlich verwendet, um sich Sprache anzueignen [Learning strategies refer to the acquisition of linguistic resources and describe the strategies that a learner uses to acquire a language with increasing autonomy and individual variation] (Niedersächsisches Kultusministerium 2015: 27).

Furthermore, the proposed lesson unit supports learning awareness ("*Lernbewusstsein*") as the students learn autonomously and discover rules by themselves: "*Die Schülerinnen und Schüler [...] finden Regeln, erkennen Fehler in Selbstkorrektur sowie peer correction und nutzen diese Erkenntnisse für den eigenen Lernprozess*" [Students [...] identify rules, recognize errors by correcting themselves or through peer correction and make use of this knowledge for their own learning process] (Niedersächsisches Kultusministerium 2015: 28).

Outline and objectives

This **90-minute lesson** is designed to be a stand-alone grammar lesson that can be embedded into any topic in which the two verbs *tell* and *say* occur, or after difficulties in the correct usage of the verbs' grammatical features arose. It has been designed with **sixth-grade students** in mind who have already been introduced to both verbs independently and are now learning about their structural differences. Since it is assumed that this is the students' first time working with a corpus, the worksheet features **step-by-step instructions** on how to extract the relevant information from the corpus. Besides providing the students with the tasks and instructions, the worksheet additionally secures all results in one document that the students can look at in case of uncertainties.

Learning objectives:

- Students understand that *tell* and *say* have different grammatical structures whose features cannot be applied synonymously.
- Students can use both verbs productively in idiomatic contexts.

This will be achieved by analysing and comparing the grammatical structures based on corpus data. In addition:

- Students develop 'corpus skills' as a learning strategy for autonomous language learning.
- Students train their communicative and social competences, e.g., by comparing and discussing their results with a partner.
- Students reflect on the experience of working with a corpus.

2 Corpus/corpora, tools and methods

As “educational systems in European countries generally [...] follow the British English model” (Naser Eddine 2011: 8), the British National Corpus (BNC) will be used to provide the necessary information for this lesson. The corpus can be freely accessed via <https://www.english-corpora.org/bnc/>. The different uses of this corpus analysis tool include “finding the frequency of words, phrases, and collocates” and “designing authentic language teaching materials and resources” (English Corpora: Most Widely Used Online Corpora. Billions of Words of Data: free Online Access, n.d.). To find out about the grammatical differences of *tell* and *say*, the students will work with the List and Collocates functions (cf. 3.3, step-by-step Instructions).

This lesson focusses on the verbs *tell* and *say*, which are introduced to German learners of English in 5th, or at the latest in 6th grade – which is the reason why this lesson was created for sixth-grade students. Once the students know both words, they should be made aware that *tell* and *say* have different grammatical structures that cannot be used synonymously. By working with a corpus, the students will learn to extract rules from authentic language data. This data-driven learning approach teaches the students how to work with the basic analysis features of a corpus tool. In the following years, more detailed and complex corpus queries can be introduced, which will build on the abilities the students have already developed. In the

context of this project, a worksheet was developed that enables the students to work autonomously, with five tasks and corresponding step-by-step instructions. It can be found at the end of this chapter.

3 Step-by-step guide

3.1 Teacher preparation

1. Become familiar with <https://www.english-corpora.org/>. Work through the worksheet yourself to be able to help your students.
2. In case you think there are too many or too few tasks, you can adjust the number according to the group of learners.
3. Does the task level match your students' skills, or might they need assistance? This is the time to think about scaffolding. Here are some ideas:
 - You could hang up a poster with the word classes on it. This could help your students in case they are not familiar with the word classes (cf. worksheet, task 3).
 - You could prepare cards containing helpful phrases (see worksheet, task 5) or beginnings of sentences (see worksheet, additional task) that the students can get from you in case they need support.
4. For a successful introduction into the lesson, draw a connection to the previous one(s), so that the students know what will happen and why. You could think about questions like:
 - Did the students have problems distinguishing *tell* and *say*?
 - Did the students ask you to explain the differences between the two verbs?
5. The conclusion period should be used to compare the students' results to see if there are any questions left and to draw a conclusion from the lesson.
 - What do I want my students to learn? What is most important to remember?
 - You could reflect on the lesson together with your students. Did they like working with a corpus autonomously? What kind of problems did they encounter?
 - Explaining how using a corpus could improve their future learning processes.
6. Add your thoughts and ideas to the provided lesson plan (cf. 3.2, lesson plan) to adjust this lesson introduction and conclusion to individual learner needs.
7. Especially if this is the first time that your students are going to work with a corpus, you should briefly introduce them to what a corpus is and what can be done with it beforehand, e.g., in the preceding lesson. That way, you can avoid overwhelming them with too much information at the beginning of this practical corpus-based lesson.

3.2 Lesson Plan

Table 1: Lesson plan

	Time	Content	Notes
Introduction	5 – 10 min	<p>Introduction of the topic and its relevance to the students</p> <p>Recap of the brief introduction to corpora</p> <p>Introduction of the lesson's outline</p>	<p>For a smooth transition, think about questions like: Why is this lesson implemented? Have the students confused the grammatical structures of <i>tell</i> and <i>say</i> in the past? Briefly explain this to your students.</p> <p>Recap: What is a corpus? What can be found out using a corpus?</p> <p>Let the students know what is going to happen during this lesson. What will they have to do? How much time do they have to work on it?</p>
Main part	60 min	<p>Teacher hands out the worksheet and answers upcoming questions.</p> <p>Students work autonomously with the worksheet</p>	<p>This does not mean walking them through the step-by-step instructions but rather clarifying comprehension questions.</p> <p>Teacher should act as learning assistant.</p> <p>If students need help, point to the scaffolding options they can use. You should try to avoid walking them through the task; let them help each other and support them in working on their own.</p> <p>Have an additional task in store for the quickest students: They could either be experts and help their classmates or get an additional task to work on.</p>
Conclusion	20 – 25 min	<p>Comparing results</p> <p>Answering remaining questions</p> <p>Reflecting on the working process</p>	<p>Compare the results and internalize the correct usage of the two verbs.</p> <p>What are the most important things you want your students to remember? e.g., rule box on the worksheet</p> <p>Short reflection to find out if there were difficulties using the corpus.</p> <p>Explain why using corpora is beneficial for the students' learning process both in and outside school.</p>

3.3 Step-by-Step Instructions

In case you are working with a corpus for the first time yourself, the following step-by-step instructions will walk you through everything you need to know for this lesson. The instructions are in chronological order based on the attached worksheet. Furthermore, there are a few additional instructions (marked with an *) to extend your basic knowledge of this web interface. This might be helpful if you want to expand the work with corpora, or if you search for different words and find out that another function might be more beneficial to reach the learning goal.

Selecting a Corpus

Open the website <https://www.english-corpora.org> and select the British National Corpus (BNC) (Fig. 1).

*You will find a listing of different corpora on the opening page. Should you be working with this analysis tool again, pick the one that is most suitable for your lesson.

Hansard Corpus	1.6 billion	British	1803-2005	Parliament
Early English Books Online	755 million	British	1470s-1690s	(Various)
Corpus of US Supreme Court Opinions	130 million	American	1790s-present	Legal opinions
TIME Magazine Corpus	100 million	American	1923-2006	Magazine
British National Corpus (BNC) *	100 million	British	1980s-1993	Balanced
Strathy Corpus (Canada)	50 million	Canadian	1970s-2000s	Balanced
CORE Corpus	50 million	6 countries	2014	Web

Fig. 1: Where to find the BNC on the homepage

If you have an account, you can log in under “My account” and “Log in”. If you do not have an account yet, you will find the “Register” button on top (Fig. 2).

The screenshot shows the homepage of English-Corpora.org. The navigation bar includes links for 'home', 'corpora', 'users', 'related resources', 'my account', 'upgrade', and 'help'. A dropdown menu is open under 'my account', showing options: 'Register', 'Log in', 'Log out', 'Name of university', 'Reset password', and 'Delete account'. The 'Log in' option is highlighted with a red box. Below the navigation bar, there is a table listing various corpora with columns for 'Corpus (online access)', 'Download', '# words', 'Dia', and 'Genre(s)'. The table lists corpora such as 'iWeb: The Intelligent Web-based Corpus', 'News on the Web (NOW)', 'Global Web-Based English (GloWbE)', and 'Wikipedia Corpus'.

Fig. 2: Logging in to english-corpora.org

Looking for Frequency

Now, you will find yourself on the SEARCH interface. On this screen, insert the word/words you are looking for. When you click on the grey words below the search box, another box with further information will pop up on the right-hand side (Fig. 3 and 4).

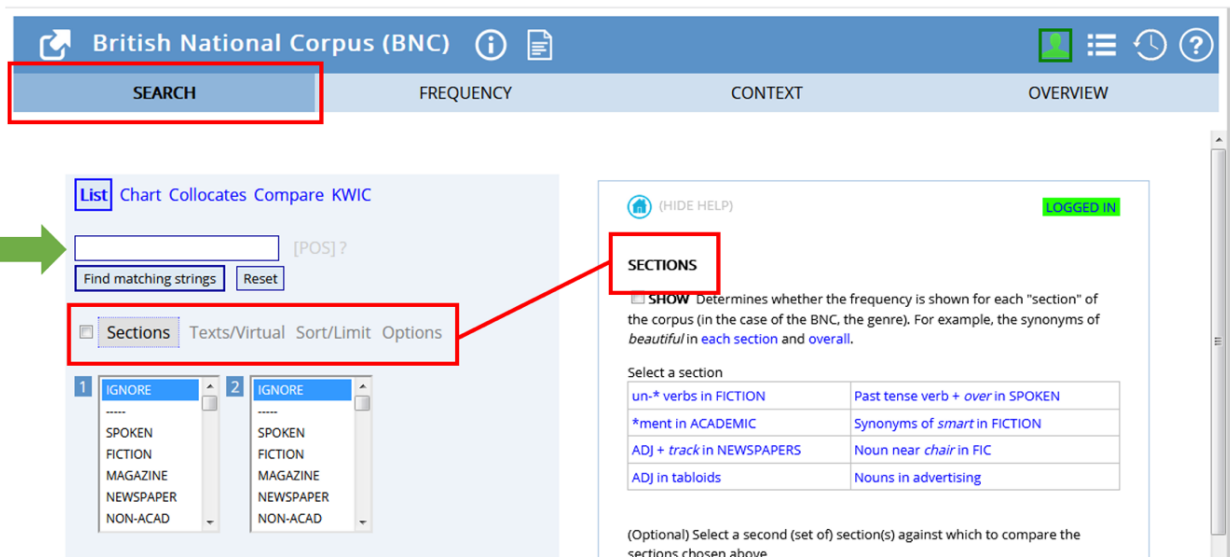


Fig. 3: The SEARCH screen and functions

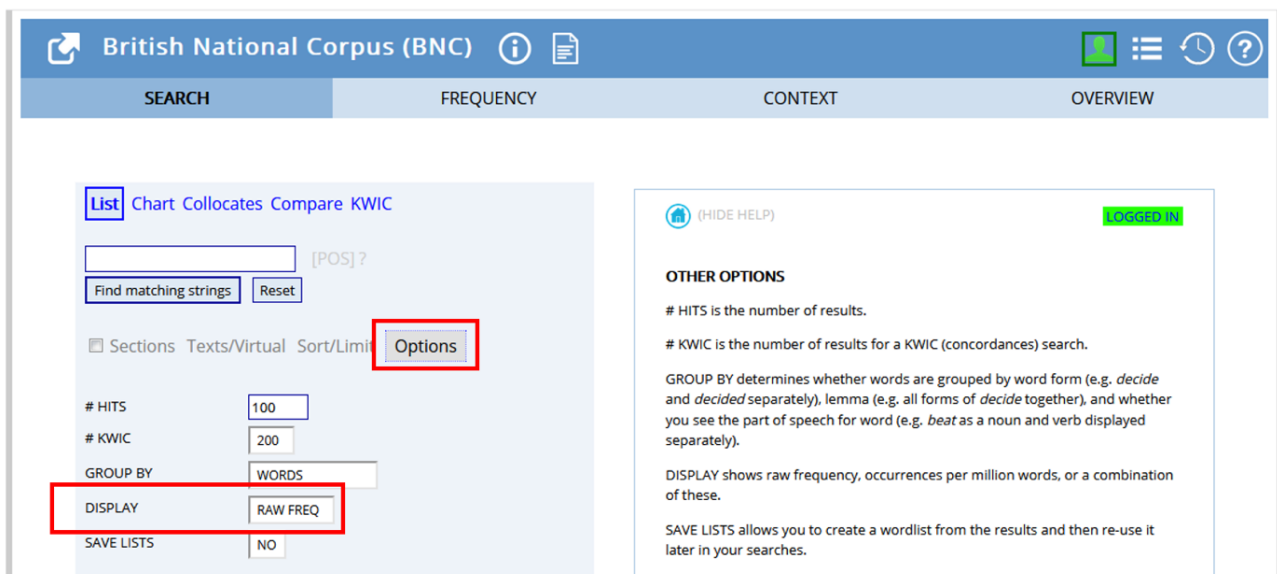


Fig. 4: How to display the number of results into frequency per million

To find out how often a word is used, simply enter the word into the white box and click on “Find matching strings”. The website will forward you to the FREQUENCY screen (Fig. 5).

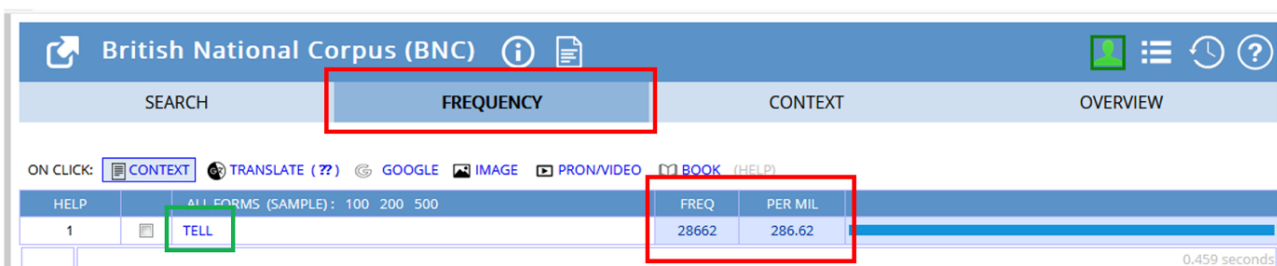


Fig. 5: The FREQUENCY screen

This interface displays the results of the search in the column “FREQ” and the column “PER MIL”. When you click on the word “TELL” (Fig. 5), the website jumps to the CONTEXT interface and provides a full list of the sentences that your search results can be found in (the concordance lines).

Lemmas

The frequency search only shows results that match the exact word form that has been written. In order to get results that contain all variants of the words (e.g., different tense and grammatical forms) the **lemma** search will be the best option. To do so, you simply enter the specific words in capital letters on the SEARCH interface (Fig. 6).

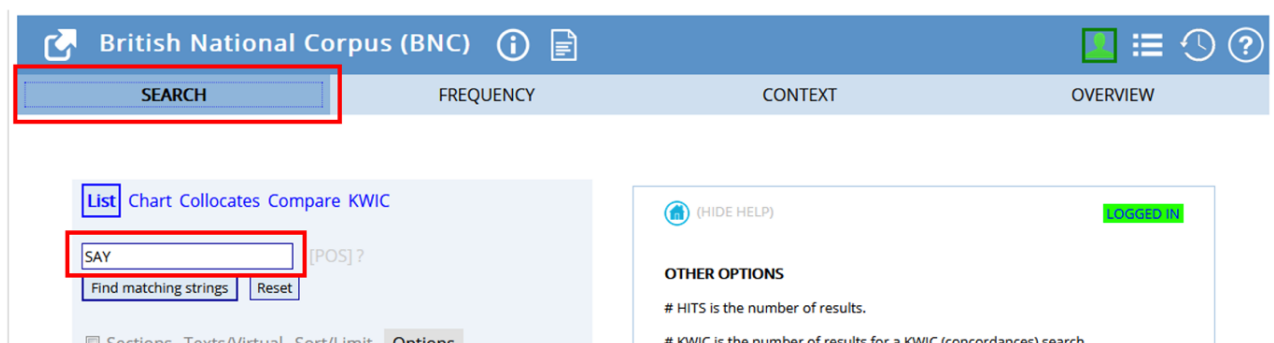


Fig. 6: The lemma search function

By confirming your entry by clicking on “Find matching strings”, you will be forwarded to the FREQUENCY interface, which displays the different variants in descending order (Fig. 7).

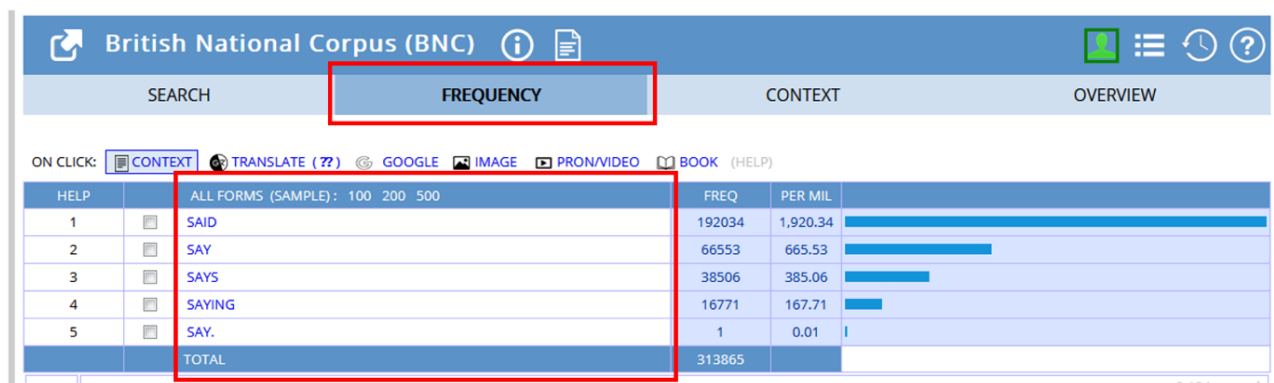


Fig. 7: Results of the lemma search

Collocations

Click on “SEARCH” in the top left corner to go back to the opening page. Otherwise, repeat “selecting a corpus”. To find out which words appear alongside *say* and *tell*, click on the Collocates function (Fig. 8).

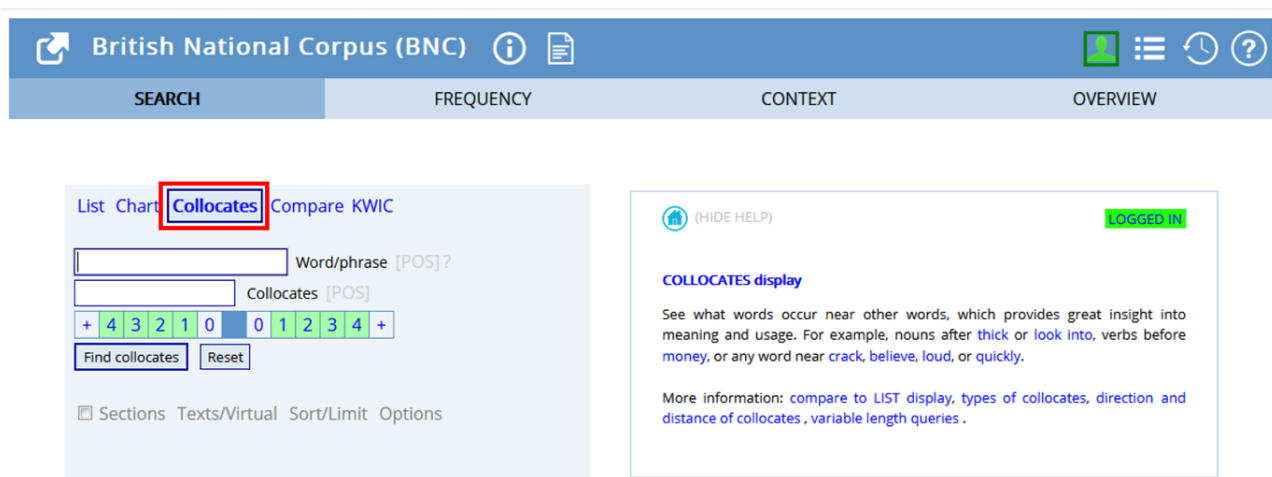


Fig. 8: The Collocates function

The task on the worksheet aims to find out how the two verbs occur with *to*. Therefore, *tell* (or *say*) needs to be written into the “Word/phrase” box and *to* into the “Collocates” box. In addition, you will also need to choose a number on the scale which indicates the positions in which the collocate appears in relation to the word. If you click on a number on the left-hand side of the numeric scale, you will get sentences that show the word *to* in front of *tell*. Clicking on a number on the right-hand side, you will get results where *to* occurs after *tell*. The given task requires to click on “3” to get results in which *to* appears within the following three words of *tell* (Fig. 9).



Fig. 9: Using the Collocates function

After clicking on “Find collocates”, the website forwards you to the FREQUENCY screen. This does not help much, as the task asks for examples. However, clicking on “TO” (Fig. 10) solves the problem and brings you straight to the CONTEXT page and to the concordance lines.

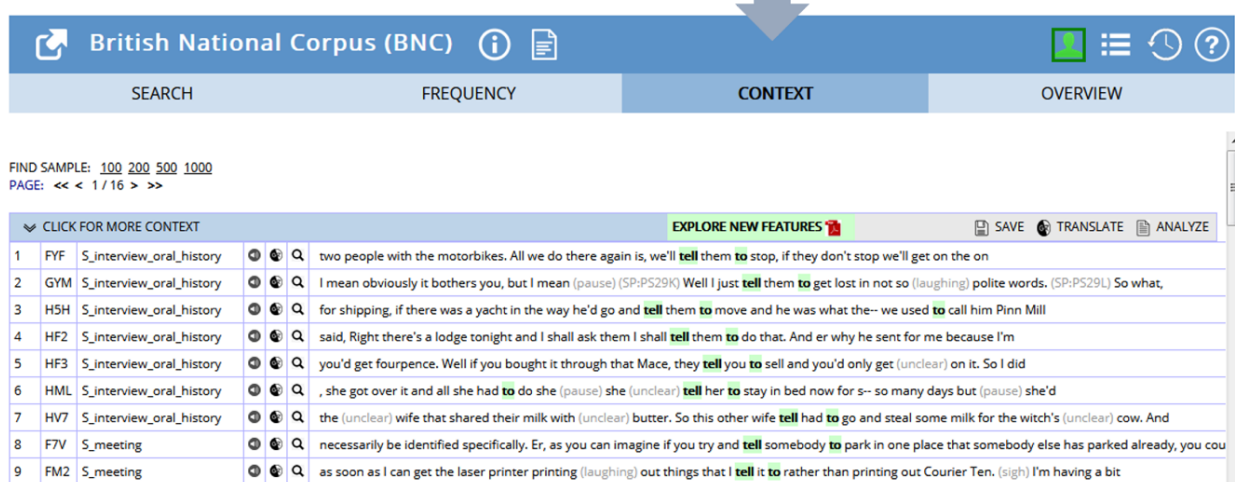
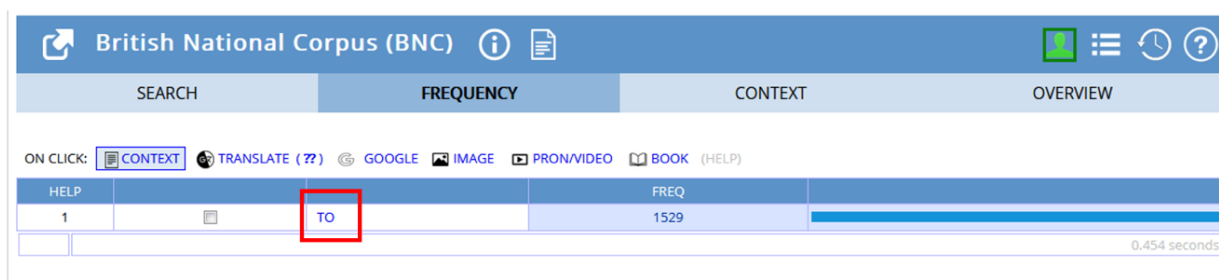


Fig. 10: How to get to the concordance lines

*The frequency interface is helpful if you want to find out which words collocate with the word/phrase in the search box. If you conduct a search like this, the FREQUENCY interface provides you with a table containing a list of the collocating words (with the most frequent at the top).

Task(s)

You will find the tasks on the worksheets attached (cf. appendix). There is a template for the students as well as a teacher's version with solutions. As mentioned before, the teacher knows their group of learners best and should adjust the scaffolding accordingly.

Follow-up activities

As this lesson focusses on grammatical difficulties in a stand-alone manner, it can be embedded into several units. The decision when to implement it must be made by the teacher. If the students have major problems with the use of this corpus analysis tool, it might be wise to divide the lesson up. The additional task is meant for students who have finished early, or could alternatively be used as homework. To enable the students to work with corpora in a more comprehensive way, the teacher should introduce them to further analysis functions from the website when the time is right, e.g., finding the correct preposition for certain words. Especially grammar and vocabulary work – which are so closely interlinked that they can hardly be addressed in isolation – can benefit from working with corpora and should therefore be regularly implemented in English classes.

4 Options and further ideas

Many words in the English language have similar meanings but are used in different contexts and/or have different grammatical structures. Native speakers intuitively learn in which contexts to use these words idiomatically. This unit was designed to teach the verbs *tell* and *say* but can also serve as a model for other words that have similar meaning, e.g., *cause* and *bring about*, *beautiful* and *attractive*, *safety* and *security*.

Another idea would be using more than one corpus to compare the use of words in different language varieties, e.g., British English (e.g., using the BNC) and American English (e.g., using the COCA). Comparisons between an inner-circle variety (e.g., British English) and an outer- or expanding circle variety (e.g., Indian or Nigerian English) are also possible. There just needs to be an appropriate corpus to use (e.g., the Global Web-Based English Corpus).

The corpus analysis tool that has been introduced in this chapter offers other search functions alongside List and Collocates, e.g., Chart and Compare. Just to give an example, the Compare function can be used to directly compare the search results of two words (Fig. 11).

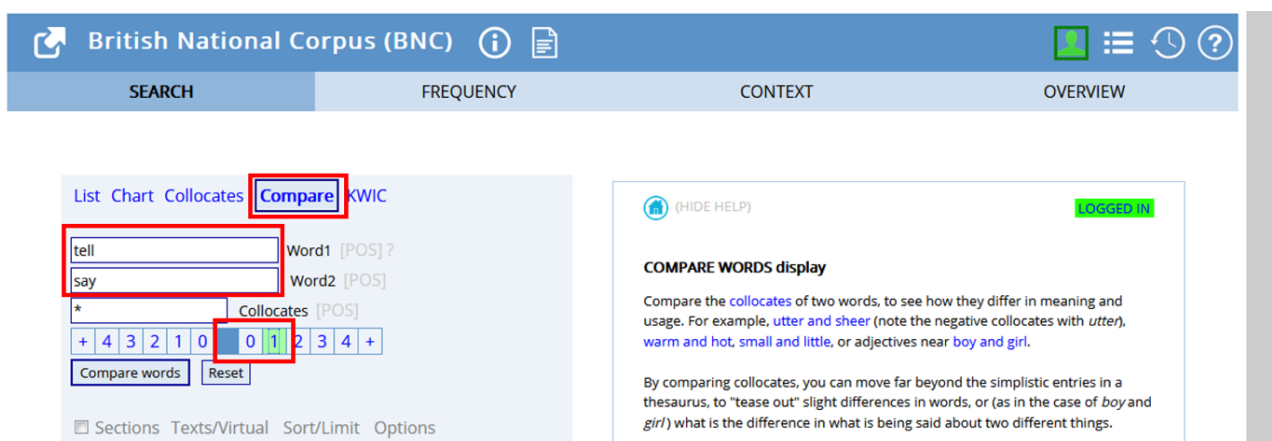


Fig. 11: The Compare function

The results are then presented in two juxtaposed tables that allow for a relatively easy comparison of their collocational patterns. The numerical scale underneath the text fields can be used to narrow the results down to the number of words that precede or follow the words. However, this function is not included in this worksheet because the simple comparison of the verbs *tell* and *say* generates the following results (Fig. 12):

WORD 1 (W1): TELL (0.43)						WORD 2 (W2): SAY (2.32)					
	WORD	W1	W2	W1/W2	SCORE		WORD	W2	W1	W2/W1	SCORE
1	ME	5216	14	372.6	865.0	1	GOODBYE	255	0	510.0	219.7
2	US	1966	6	327.7	760.8	2	WELL	457	1	457.0	196.8
3	HIM	2180	8	272.5	632.7	3	HELLO	205	0	410.0	176.6
4	LIES	58	0	116.0	269.3	4	YES	296	1	296.0	127.5
5	STORIES	51	0	102.0	236.8	5	THANK	124	0	248.0	106.8
6	ME.	45	0	90.0	209.0	6	SO	649	3	216.3	93.2
7	HERSELF	35	0	70.0	162.5	7	ERM	92	0	184.0	79.2
8	ANYBODY	69	1	69.0	160.2	8	LOOK	85	0	170.0	73.2
9	MYSELF	58	1	58.0	134.7	9	DO	76	0	152.0	65.5

Fig 12: Compare function results for 'say' and 'tell'

Abstracting results from this chart did not seem age-appropriate to us and would probably leave students confused. Should the structure of this worksheet be copied and adapted for older students (presumably grade 8 and higher), including the Compare function could be considered as a way of exposing the students to additional input that they may benefit from.

5 Conclusion

This lesson introduces students to the work with corpora at an early stage. Acquiring digital skills is of importance in these times of rapid technical developments. Consequently, it may be beneficial to introduce students to a variety of digital tools. The more they have the opportunity to interact with corpus platforms, the easier the handling gets. This lesson could therefore be useful to introduce students to basic corpus analyses. It provides a lot of instructions and scaffolding which supports those in need, especially when using this type of tool for the very first time.

Language is not a fixed concept; therefore, it is relevant to use tools that allow the student to dive into real language situations aside from the textbook used in school. Our lesson focusses on a typical learner error, namely the collocational patterns of the verbs *tell* and *say*. German-speaking students cannot be expected to understand why they cannot use the words synonymously without further explanations. Implementing this lesson to clarify the distinction at an early stage can avoid long-term fossilisation of unidiomatic structures.

By following the step-by-step instructions included in this chapter, teachers should be able to create corpus-based lessons on any other typical error that their students regularly make.

6 Resources and references

Davies, Mark. 2020. English Corpora. <https://www.english-corpora.org/> (25 March, 2021).

English Corpora: Most widely used online corpora. Billions of words of data: free online access. n.d. <https://www.english-corpora.org/overview.asp> (12 August, 2020).

Naser Eddine, A. 2011. Second language acquisition: The articulation of vowels and the importance of tools in the learning process. In J. Arabski & A. Wojtaszek (eds.), *The acquisition of L2 phonology*, 3-15. Multilingual Matters: Bristol.

Niedersächsisches Kultusministerium. 2015. Kerncurriculum für das Gymnasium, Schuljahrgänge 5-10. https://cuvo.nibis.de/cuvo.php?skey_lev0_0=Schulbereich&svalue_lev0_0=Sek+I&skey_lev0_1=Fach&svalue_lev0_1=Englisch&skey_lev0_2=Dokumentenart&svalue_lev0_2=Kerncurriculum&docid=1197&p=detail_view (26 March, 2021).

Turton, N. D., & J. B. Heaton. 1997. *Longman dictionary of common errors* (3. impr). Harlow: Longman.

7 Appendix

7.1 Worksheet “Tell or Say” (four pages) – Student version

Tell or Say Student worksheet

1. Fill in the gaps

Fill in the gaps with *tell* or *say*.

Keep in mind: Sometimes you'll have to use the past form.

- She _____ that she went to the cinema yesterday.
 He _____ the teacher that he was at school yesterday.
 I have to _____ you something!
 What did she _____?
 What did Max _____ the teacher?
 The girl on the bus _____, "Good morning, Julie!"
 Laura _____ me that she played tennis last week.

2. Say to me & tell to me?

You're not sure which combination to use? Let's find out!

To do so, follow the instructions below:

1. Open the website www.english-corpora.org/bnc/.
2. There is a blue box on the left-hand side of the page. Select the word **List** by clicking on it.
3. Type **SAY TO ME** in the white box.
4. Click on **Find matching strings** below the white box in which you typed the words.
5. The search results are presented in a table. There is a column⁴ labeled **FREQ** (frequency). This number tells you how often *say to me* was found in the corpus. Write the total of all numbers into the table below.
6. Repeat the steps 1-6 with the phrase **TELL TO ME**.

Not sure how to do it?
Ask a friend or me. ☺

Frequency	
Say to me	
Tell to me	

What do the results tell you about the two combinations? Are both of them used by British speakers of English?

This worksheet was downloaded from: Dürnhaupt, Geeske & Lena Haase (2021). "Differences in the grammatical structures of *tell* and *say*." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

3. When to use *to*

When can you use *to* with *tell* or *say*? Use the following instructions to find out yourself:

1. Open the website www.english-corpora.org/bnc/ again. If that website is still open, you can click on **SEARCH** in the top left corner.
2. This time select the **Collocates** option in the blue box on the left-hand side. This tells you which words appear together.
3. Type *tell* in the white box labelled "word/phrase" and type *to* in the white box labelled "collocates".
There is a scale with numbers beneath. Click on the **3 to the right** of the middle. It should look like this:



4. This means we can look at sentences where *to* appears together with *tell*, at the most three words after *tell*.
5. Good job, you have set up the search! Click **Find collocates** beneath the scale.
6. Now, you can see the results in a table. Click on the word **TO**.
7. Here is a list of all the sentences that match your search.

Write down the three words after *tell* for sentences 1 to 6. Colour *tell* and *to* in a colour of your choice.

- 1) tell _____ ... 2) tell _____ ... 3) tell _____ ...
4) tell _____ ... 5) tell _____ ... 6) tell _____ ...

What type of word comes right after *tell*?

Do you need help?
Come and see me. 😊

What type of word comes right after *to*?

Repeat steps 1-6 with *say*.

Write down the three words after *say* for sentences 1 to 6. Colour *say* and *to* in a colour of your choice.

- 1) say _____ ... 2) say _____ ... 3) say _____ ...
4) say _____ ... 5) say _____ ... 6) say _____ ...

Do you notice the difference? Fill in gaps and compare with a partner.

Do you have different answers?
Come and see me. 😊

This worksheet was downloaded from: Dünhaupt, Geske & Lena Haase (2021). "Differences in the grammatical structures of *tell* and *say*." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

Rules

_____ and _____ are used in different ways. Here is what you can say:

→ _____ *someone something* or _____ *someone to do something*

→ _____ *(something) to someone*

Note: The word *to* is not used directly after _____.

Congratulations, you are an expert now!

4. Correct or incorrect?

Are the following sentences correct or incorrect?

If they are incorrect, underline the problematic word and correct the sentence.

Note: You have to use the simple past.

You are not sure?
Look at the rule box.

1) "Good morning," she told to me.

2) Lisa asked me if I liked chocolate and I told yes.

3) Ben told me about his new basketball.

4) He said, "I really like the sunny weather."

5) Dad told to me to do my homework.

6) My teacher said me that I was a good student.

This worksheet was downloaded from: Dürnhaupt, Geske & Lena Haase (2021). "Differences in the grammatical structures of *tell* and *say*." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

5. Think – Pair – Share

Think: Think about two things you did last weekend. You may take some notes if that helps you.

Pair: Tell your partner about the activities you did last weekend and write down what your partner did.

Share: With the whole class: Tell us what your partner did last weekend. Use the words *tell* and *say*.

You are not sure how to build those sentences?
I can help. 😊

Additional Task

Get creative: Write a short story (five to ten sentences) including *tell* and *say*.

This worksheet was downloaded from: Dürnhaupt, Geske & Lena Haase (2021). "Differences in the grammatical structures of *tell* and *say*." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

7.2 Worksheet “Tell or Say” (four pages) – Teachers’ version

Tell or Say Teachers' version

1. Fill in the gaps

Fill in the gaps with *tell* or *say*.

Keep in mind: Sometimes you have to use the past form.

She **said** that she went to the cinema yesterday.

He **told** the teacher that he was at school yesterday.

I have to **tell** you something!

What did she **say**?

What did Max **tell** the teacher?

The girl on the bus **said**, "Good morning, Julie!"

Laura **told** me that she played Tennis last week.

2. Say to me & tell to me

You're not sure which combination to use? Let's find out!

To do so, follow the instructions below:

1. Open the website www.english-corpora.org/bnc/.
2. There is a blue box on the left-hand side of the page. Select the word **List** by clicking it.
3. Type **SAY TO ME** in the white box.
4. Click **Find matching strings** beneath the white box in which you typed the words.
5. The search results are presented in a table. There is a column labeled **FREQ** (frequency). This number tells you how often say to me was found in the corpus. Write the total of all numbers into the table below.
6. Repeat the steps 1-6 with the word **TELL TO ME**.


	Frequency
Say to me	1141
Tell to me	12

What do the results tell you about the two combinations? Are both of them used?

Tell to me is not used as often as *say to me* → *say to me* is very rarely used

3. When to use *to*

Now: When can you use *to* with *tell* or *say*? Use the following instructions to find out yourself:

1. Open the website www.english-corpora.org/bnc/ again. If that website is still open, you can click on **SEARCH** in the top left corner.
2. This time select the word **Collocates** in the blue box on the left-hand side. This tells you which words appear together.
3. Type **tell** in the white box labelled "word/phrase" and type **to** in the white box labelled "collocates".
4. There is a scale with numbers beneath. Click on the **3** to the right of the middle. It should look like this:


This means we can look at sentences where *to* appears together with *tell*, at the most three words after *tell*.

5. Good job, you have set up the search! Click **Find collocates** beneath the scale.
6. Now, you can see the results in a table. Click on the word **TO**.
7. Here is a list of all the sentences that match your search. These sentences are called concordance lines.

Write down the three words after *tell* for sentences 1 to 6. Colour *tell* and *to* in a colour of your choice.

- 1) **tell** them **to** stop... 2) **tell** them **to** get... 3) **tell** them **to** move...
4) **tell** them **to** do ... 5) **tell** you **to** sell... 6) **tell** her **to** stay ...

What type of word comes right after *tell*?

Personal pronoun

What type of word comes right after *to*?

Verb (infinitive)

Repeat steps 1-6 with *say*.

Write down the three words after *say* for sentences 1 to 6. Colour *say* and *to* in a colour of your choice.

- 1) **say** to you know... 2) **say** to show and... 3) **say** to us. And...
4) **say** to Alan Yeah... 5) **say** to the County Council... 6) **say** to them well ...

Do you notice the difference? Fill in gaps and compare with a partner.

This worksheet was downloaded from: Dürnhaupt, Geske & Lena Haase (2021). "Differences in the grammatical structures of *tell* and *say*." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

Rules

Tell and *say* are used in different ways. Here is what you can say:

→ *Tell* someone something or *tell* someone to do something

→ *Say* (something) to someone

Note: The word *to* is not used directly after *tell*.

4. Correct or incorrect?

Are the following sentences correct or incorrect?
If they are incorrect, underline the problematic word and correct it.
Note: You have to use the simple past.

- 1) "Good morning," she **told** to me.

"Good morning," she said to me.

- 2) Lisa asked me if I liked chocolate and I **told** yes.

Lisa asked me if I liked chocolate and I said yes.

- 3) Ben told me about his new basketball.

correct

- 4) He said, "I really like the sunny weather."

correct

- 5) Dad **told to** me to do my homework.

Dad told me to do my homework.

- 6) My teacher **said** me that I was a good student.

My teacher told me (that) I was a good student.

5. Think – Pair – Share

Think: Think about two things you did last weekend. You may take some notes if that helps you.

Pair: Tell your partner about the activities you did last weekend.

Share: Your partner has to report what you did last weekend with the help of *tell* and *say*. If you need more help with building such sentences, you may come to see me.

Individual answers

This worksheet was downloaded from: Dürnhaupt, Geske & Lena Haase (2021). "Differences in the grammatical structures of *tell* and *say*." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

Additional Task

Get creative: Write a short story (five to ten sentences) including *tell* and *say*.

Individual answers

This worksheet was downloaded from: Dürnhaupt, Geske & Lena Haase (2021): "Differences in the grammatical structures of *tell* and *say*." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefall.pressbooks.com>. CC-BY-NC 4.0.

1 For a very brief introduction to Braj Kachru's model of World Englishes, see: https://en.wikipedia.org/wiki/World_Englishes#Kachru's_Three_Circles_of_English (25 March, 2021).

5. Using corpora in the EFL classroom for creative writing

The case of descriptive adjectives

JONAS DAMMANN

1 Introduction

“In SLA [Second Language Acquisition] literature, there is general agreement that all language acquisition starts with input” (Perez et al 2014: 173). In other words, students learning English as a Foreign Language (EFL) need to be confronted with language (input) in order to be able to produce language (output). Writing tasks target students’ productive skills. However, it is simply impossible to expect students to use vocabulary which they have never been exposed to before. Therefore, they need to be given opportunities to encounter such suitable vocabulary beforehand. This chapter thus aims to show how English corpora can be used to provide such input.

The English curriculum for secondary schools in Lower Saxony explicitly lists “creative writing” as a competence (Niedersächsisches Kultusministerium 2015: 20). By the end of 8th grade (end of middle school), students are expected to be able to produce simple essays. By the end of 10th grade, this competence is to be improved: students are expected to write more complex and extensive texts. Based on these curricular guidelines, the exercises in this chapter focus on students’ creative writing skills.

Outline and objectives

- The activities presented in this chapter are mainly designed for capacity building. This chapter will show how students can be taught to improve their creative writing skills autonomously by consulting corpora. The exercises provide valuable input in the form of suitable vocabulary, and also teach how to find synonyms to expand and diversify students’ vocabulary. The example in this chapter concentrates on writing a travel blog post. However, the tasks can easily be adapted to different writing tasks, too.
- This chapter targets 8th grade students at secondary school. The lesson plan is designed for a 90-minute double lesson. Introducing the use of corpora as a writing resource as early as possible is recommended as students will benefit from that point on. At this age, students are expected to only have quite a limited vocabulary range, hence, such a strategy to expand their vocabulary can be helpful. Furthermore, 8th grade students can reasonably be expected to be technically able to operate corpus analysis tools without any great difficulties.

2 Tools and methods

This chapter relies on Sketch Engine as an online corpus analysis tool. This tool already contains more than 500 ready-to-use corpora in different languages and also makes it easy for users to optionally compile their own corpora. It is accessed via its website, therefore no additional program or application must be downloaded; a working internet connection suffices. Although Sketch Engine is not free to use, a number of reasons justify the choice of this tool for the present chapter. First of all, it provides powerful inbuilt corpus analysis functions which can be used to analyze collocations, word lists, concordances, and much more. Sketch Engine comes with a free 30-days-trial and is also free until March 2022 for researchers, lecturers and students of EU academic institutions taking part in the ELEXIS project.

The corpus used in this chapter is the English Web 2015 (abbreviated to enTenTen15), which consists of ca. 15.7 billion words of written language from the internet. The English Web 2015 corpus was chosen because it is one of the largest corpora available and, since it is a compilation of online texts, is well suited for the exercise of writing a travel blog post.

The lesson plan in 4.2. presents the important functions of the Sketch Engine interface to students. During the second half of the lesson, students can then make use of their new knowledge and consult a corpus in order to find suitable vocabulary for a short essay. The lesson plan may be embedded in an introductory lesson on corpus methods or immediately afterwards. However, it is not advised to start using Sketch Engine without any introduction, as this will likely overwhelm students.

3 Step-by-step guide

Note that, for the lesson presented, each student or pair of students must have a laptop or computer with a working internet connection, as well as a Sketch Engine account.

1. **Introduction:** The teacher's laptop is connected to the video projector to project a short text on the board:
Last weekend I went to the beach. The beach was beautiful and the sea was also beautiful. The beach was very bright and reflected the sunlight. Along the beach there were many beautiful cafes and restaurant. I walked on the beach for many hours, just up and down. The beach was not too far from my hotel, so I stayed until it got dark. What a beautiful day.
2. The teacher asks the students what they think of the text and how it can be made more fun to read and more varied in terms of vocabulary. At this point, the discussion should lead towards the importance of adequate adjectives for more vivid descriptions and the use of synonyms to avoid repetition.
3. Students turn on their devices, connect to www.sketchengine.eu and log in. The teacher also follows these steps. As the teacher's laptop is still connected to the video projector, the students can compare their screens in order to check whether they have followed the steps correctly.
4. In the search bar at the top of the page, the students are instructed to load the "English Web 2015" corpus (Fig. 1).

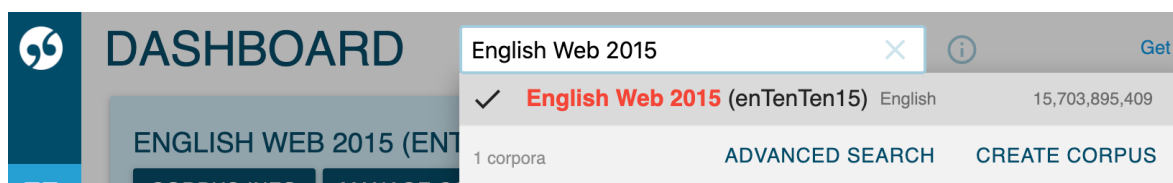


Fig. 1: Lesson plan: Step 4.

- Students are instructed to click on “Word Sketch”, a function that lists a word’s most typical collocates, organised according to their grammatical relation to the search word (Fig. 2).



Fig. 2: Lesson plan: Step 5.

- Students are asked to type in *beach* and click on “GO” (Fig. 3).

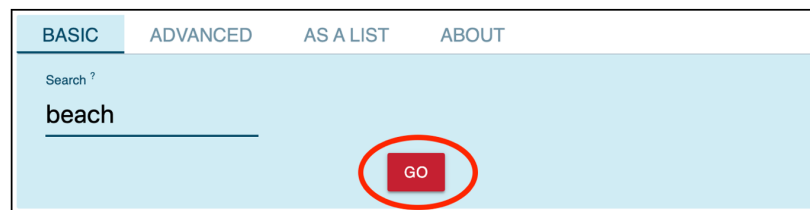


Fig. 3: Lesson plan: Step 6.

- The teacher gives a brief presentation of the different kinds of grammatical relations and their value for creative writing. In this case, the category “modifiers of ‘beach’” is the most important, as it lists all the words which are frequently used to modify “beach” in the corpus, and these are mostly adjectives. In this example, the collocation *beautiful beach* is very frequent, but alternatives could be *stunning beach* or *pristine beach* (Fig. 4).

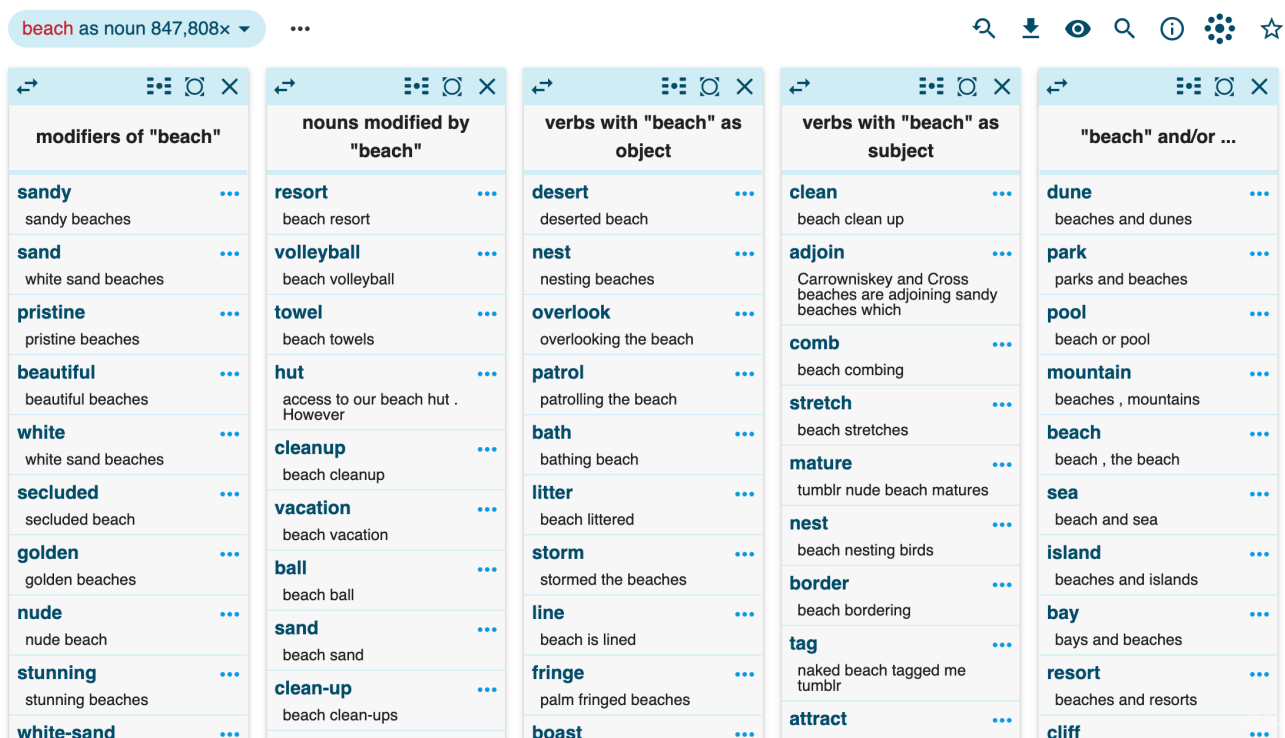


Fig. 4: Lesson plan: Step 7.

8. Students are then encouraged to use this method to find suitable adjectives to avoid the overuse of *beautiful* in the introductory text. Weaker students may need to be told that the words they will need to look up using the “Word Sketch” function to achieve this, e.g. *sea*, *cafés*, *restaurants*, and *day*. Thus, for example, for *sea*, students may choose to replace *beautiful* in the original text with *calm*, *deep*, *blue*, *crystal clear*, etc.
9. In the next step, students are introduced to the “Thesaurus” function, which generates a list of the search word’s synonyms and related words. To use this function, they must click on “Thesaurus” in the menu (Fig. 5).

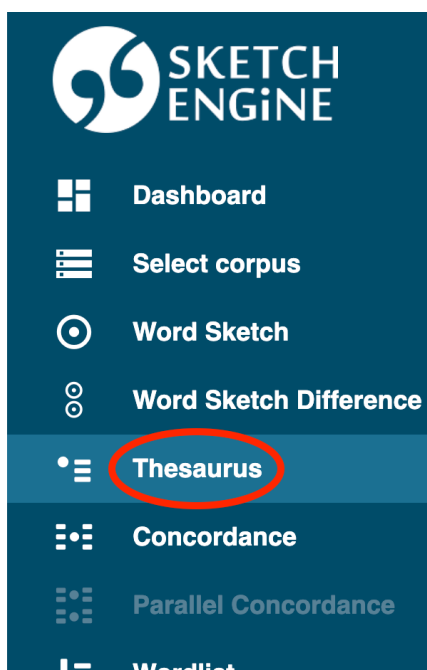


Fig. 5: Lesson plan: Step 9.

10. As an example, students are asked to search for *beautiful* (Fig. 6).

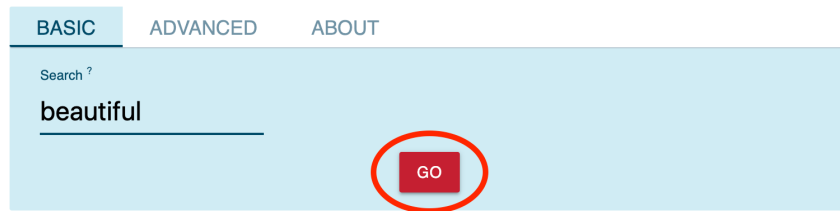


Fig. 6. Lesson plan: Step 10.

11. The teacher explains that the resulting list of synonyms was automatically compiled and may therefore contain anomalies. However, this list can nevertheless be very useful in helping students diversify their vocabulary in their own creative writing. By clicking on the three dots to the right of any word, students can see the word used in context and hence check that it is appropriate for their own text (Fig. 7).

	Word	Frequency ?		Word	Frequency ?
1	lovely	470,116 ...	11	attractive	444,497 ...
2	wonderful	1,312,241 ...	12	impressive	514,924 ...
3	amazing	1,403,466 ...	13	bright	716,579 ...
4	gorgeous	216,781 ...	14	dark	1,018,354 ...
5	stunning	313,407 ...	15	brilliant	373,851 ...
6	nice	1,485,424 ...	16	incredible	583,267 ...
7	cool	763,678 ...	17	elegant	213,207 ...
8	awesome	611,272 ...	18	fine	1,077,518 ...
9	fantastic	671,620 ...	19	spectacular	271,215 ...
10	perfect	1,551,856 ...	20	rich	1,227,301 ...

Rows per page: 20 ▼ 1–20 of 1,000

Fig. 7. Lesson plan: Step 11.

12. The teacher then provides and explains the following tasks:

A) What modifiers are frequently used in the corpus to describe a village? Use Sketch Engine's "Word Sketch" function to find these modifiers and write a short fictional paragraph about your visit to a village (ca. 5-7 sentences). Try to make your text as varied as possible.

B) Try using Sketch Engine's "Thesaurus" function to find suitable synonyms for some of these modifiers.

13. In the end, students can present their results and also the difficulties they encountered. Other students try to find alternative solutions.

4 Options and further ideas

As mentioned above, the guides and exercises presented in this chapter serve solely as examples and can easily be adapted to other learning objectives and topics. For example, students could use corpora as an aid for writing manuals, e-mails, short stories, and so on. With Sketch Engine, it is also possible to focus more specifically on certain word classes, e.g. prepositions or adverbs.

5 Conclusion

This chapter aimed to illustrate how students can improve their creative writing skills by consulting a corpus analysis tool. Students can use Sketch Engine to extend and diversify their own vocabulary. Furthermore, Sketch Engine has the potential to be used for other exercises designed to improve the students' writing skills.

However, it has to be taken into consideration that Sketch Engine is not free to use beyond the 30-day trial period. Alternatively, teachers may want to consult SKELL (skell.sketchengine.co.uk), which is a free, highly simplified, version of Sketch Engine designed specifically for language learning. Whilst it is based on a smaller corpus, incidentally also of web-based language, it nevertheless also includes the functions "Word Sketch" and "Similar words" (which is similar to the "Thesaurus" function), which can be used for exactly the same types of activities described in this chapter. As with all new tools, it is recommended that teachers test the interface beforehand to check that the results provided by SKELL are likely to meet the needs of the students.

6 Resources and references

Kilgarriff, Adam, Pavel Rychlý, Pavel Smrž & David Tugwell. 2004. The Sketch Engine: Information technology. https://www.sketchengine.eu/wp-content/uploads/The_Sketch_Engine_2004.pdf (13 October, 2020).

Niedersächsisches Kultusministerium. 2015. Kerncurriculum für das Gymnasium, Schulgänge 5-10: Englisch. <https://cuvo.nibis.de/cuvo.php?p=download&upload=139> (13 October, 2020).

Perez, Maribel Montero, Hans Paulussen, Lieve Macken & Piet Desmet. 2014. From input to output: The potential of parallel corpora for CALL. *Lang Resources & Evaluation* 48(1). 165-189. DOI: 10.1007/s10579-013-9241-4.

6. How to do data-driven learning (DDL) with secondary school pupils to improve their writing

KATHARINA BUSMANN

1 Introduction and rationale

According to Anne Wichmann et al. (1997), the main rationale for using corpora in teaching is their immediate availability for students' use. To do so, students need to acquire the essential know-how to be able to explore corpora as part of their self-regulated learning (cf. Wichmann et al. 1997: 7). Likewise, the curriculum for secondary schools in Lower Saxony requires students to acquire essential learning strategies and study techniques that enable them to solve problems effectively on their own using a range of resources (cf. Niedersächsisches Kultusministerium 2015: 27). Accordingly, this chapter focuses on how teachers can introduce data-driven learning (DDL) to their students so that they can autonomously use corpora to improve their writing after having a writing assignment marked by the teacher and returned to them. Additionally, it gives further ideas on how to use the Corpus of Contemporary American English (COCA) for writing in general, since it allows learners to quickly and easily see how native speakers use language in a wide variety of texts.

Outline and objectives

The proposed lesson...

- is suitable for secondary school pupils at the age of 12 or older.
- includes a pre-task which introduces the COCA and explains the basic functions of the web interface.
 - Preparation time: ca. 15-20 min.
 - Class time: ca. 10-15 min.
- features a main task guided by a worksheet, so that students learn to correct some of their own mistakes independently thanks to the information gleaned from the COCA.
 - Preparation time: none
 - Class time: ca. 20-25 min.
- ends with three review questions that encourage students to reflect on the use of the concordance function.
 - Preparation time: none
 - Class time: ca. 10 min.

In light of the above, the following learning objectives emerge:

After the lesson, students ...

- can search for concordance lines in the COCA.
- can communicate how specific words or phrases are used in writing by consulting the COCA.
- can correct their mistakes independently by inferring information given by the COCA and thus improve their writing skills.
- can reflect on the utility of the concordance function.

If you are familiar with using the COCA, there is hardly any teacher preparation needed for this lesson. Otherwise, please make sure that you work through the introductory steps below beforehand, in order to be able to explain the tools to your students and answer students' questions in class.

2 Corpus, tools and methods

The resource used for this lesson is the COCA (Corpus of Contemporary American English), which contains more than one billion words and is equally divided up into spoken English, fiction, popular magazines, newspapers and academic texts. With the latest update in March 2020, TV and movie subtitles, blogs, and other webpages have been added to the corpus as well (cf. Davies 2020b: 1). According to english-corpora.org, the COCA is the most widely used corpus of general English. Since the latest texts are dated from 2019, it is also the most up to date (cf. *ibid*). Once you have learnt about its basic functions, the web interface is very easy to handle. It is free for all to use. It follows from all these considerations that the COCA is a good choice of resource for such a lesson.

3 Step-by-step guide

Lesson plan

The lesson plan (Table 1) gives an overview of how the lesson may be structured.

Table 1: Lesson Plan.

	Time	Activity	Materials
Pre-Task	10-15'	<p>a) Practice searching in the COCA by working through the <i>fewer/less</i> example in plenum</p> <p>or b) Choose any other suitable warm-up activity (e.g. <i>Spot the mistake</i>: Write lexically or grammatically unidiomatic sentences on the blackboard and ask your students to find and correct the mistakes (cf. Boerger 2020: 2)).</p>	Computers, copies of the instructions
Main Task	20-25'	Students complete the worksheet by consulting the COCA	Computers, worksheet (see appendix)
Review	10'	Students answer the review questions.	Review questions

Instructions

The following instructions cover explanations about general functions of the web interface while working through an example that focuses on the different usages of two commonly confused words, namely *fewer* and *less*. They both have the same meaning. However, when do you use *less* and when should you use *fewer*? Perhaps, this confusion is a mistake which quite a few students made in the writing assignment that you recently corrected and aim to hand back in the designed lesson.

Steps:

1. Go to www.english-corpora.org.
2. Click on “Corpus of Contemporary American English (COCA)” to open the corpus (Fig. 1).

The screenshot shows the English-Corpora.org website. The header includes the site name and navigation links: home, corpora, users, related resources, my account, upgrade, help. Below the header, there is a list of corpora with columns for Corpus (online access), Download, # words, Dialect, Time period, and Genre(s). The 'Corpus of Contemporary American English (COCA)' is highlighted with a red circle.

Corpus (online access)	Download	# words	Dialect	Time period	Genre(s)
iWeb: The Intelligent Web-based Corpus	📄	14 billion	6 countries	2017	Web
News on the Web (NOW)	📄	9.4 billion+	20 countries	2010-last month	Web: News
Global Web-Based English (GloWbE)	📄	1.9 billion	20 countries	2012-13	Web (incl blogs)
Wikipedia Corpus	📄	1.9 billion	(Various)	2014	Wikipedia
Corpus of Contemporary American English (COCA)	📄	1.0 billion	American	1990-2019	Balanced
Corpus of Historical American English (COHA)	📄	400 million	American	1810-2009	Balanced
The TV Corpus	📄	325 million	6 countries	1950-2018	TV shows
The Movie Corpus	📄	200 million	6 countries	1930-2018	Movies
Corpus of American Soap Operas	📄	100 million	American	2001-2012	TV shows

Fig. 1: Selecting a corpus.

3. You should now be able to see the following display (Fig. 2):

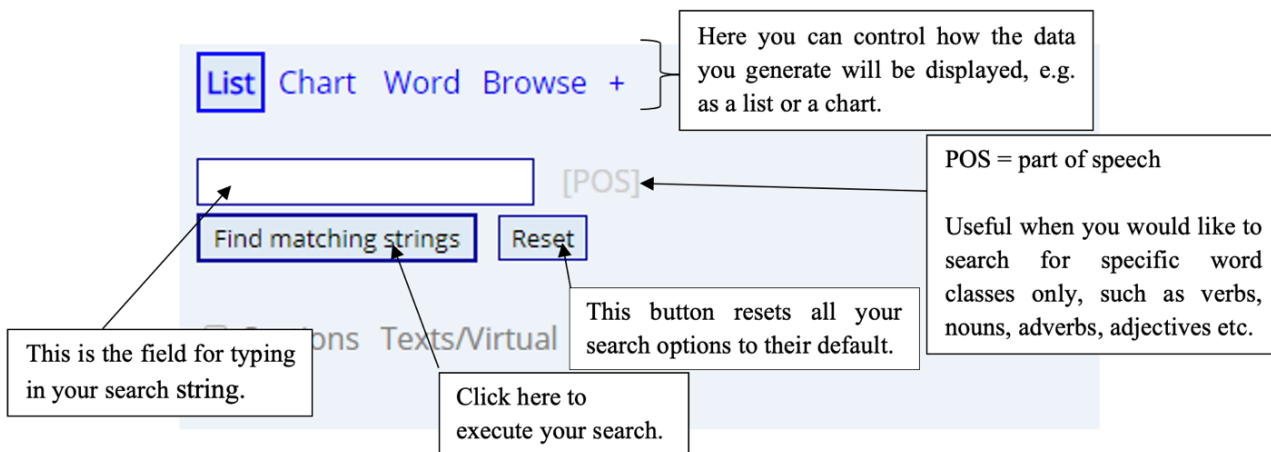


Fig. 2: Conducting a search in the COCA.

4. Click on “Sections” (Fig. 3).

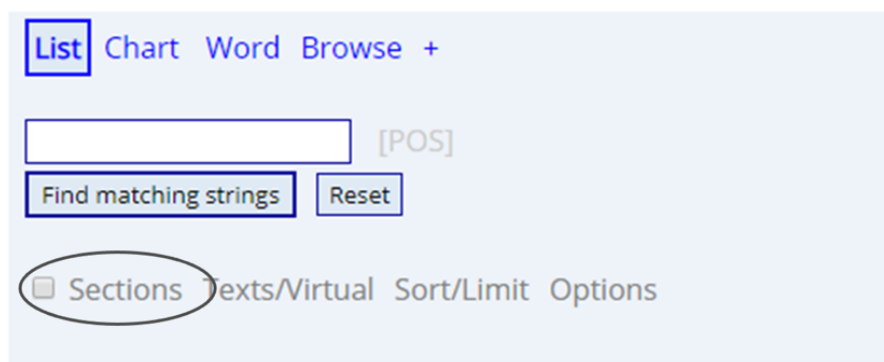


Fig. 3: Conducting a search in the COCA.

5. Choose the register(s) you want to focus on. If you want to find out how *less* and *fewer* are used in typical written text types, select “Fiction”, “Magazine”, “Newspaper” and “Academic”, for instance (Fig. 4).

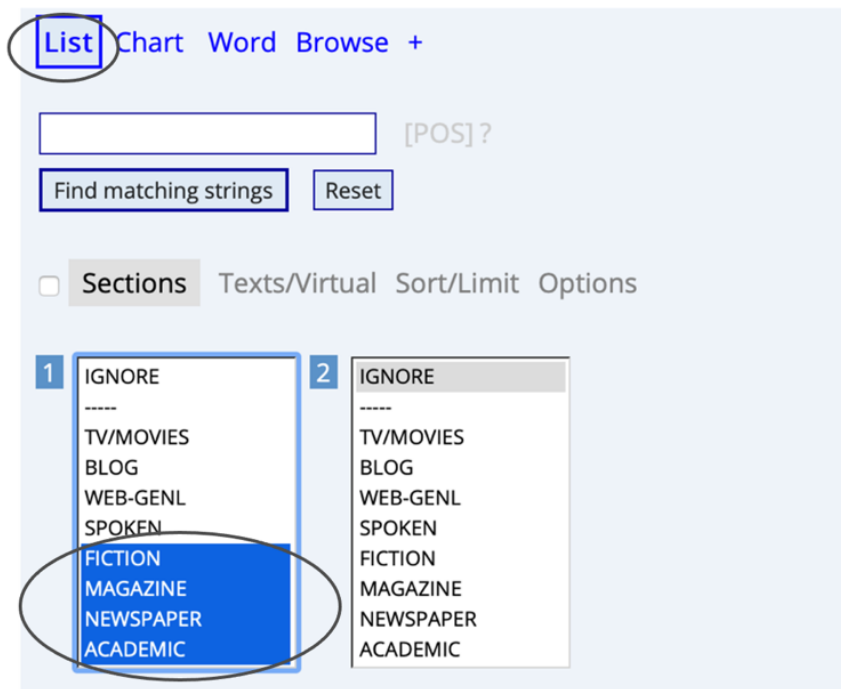


Fig. 4: Limiting the search to specific registers.

6. Make sure you have selected the “List”-option. Then, type in *fewer* and press “Find matching strings”.
7. Click on the word to see in which contexts it has been used (Fig. 5).

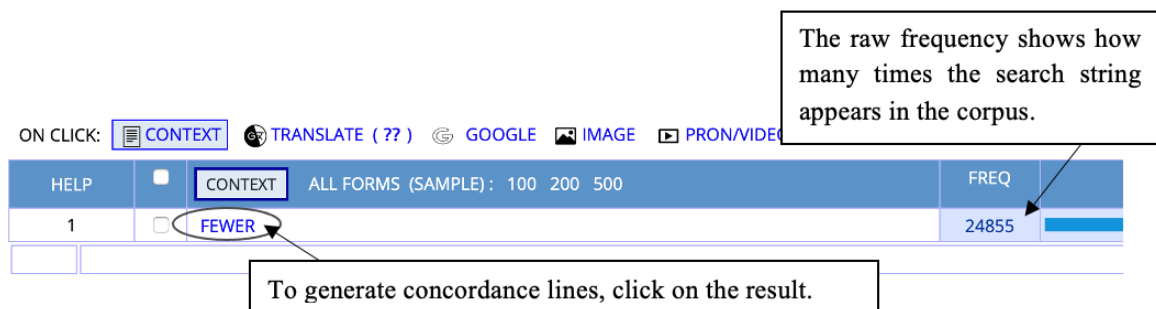


Fig. 5: Result page from the corpus query.

8. In the concordance lines you can now look for regularities in the usage of the word *fewer* (Fig. 6).

Here you can draw a random sample of concordances and control the number of results displayed. Each time you click on '100', for instance, you get a new random set of 100 sample concordance lines.

FIND SAMPLE: [100](#) [200](#) [500](#) [1000](#)
 PAGE: << 1 / 249 >>

CLICK FOR MORE CONTEXT [?] [?]

1	1990	MAG	AmerArtist	A B C	when the demand is great and is able to work more slowly when he has fewer customers. Christmas has proven to be a very busy season; however, his	
2	2006	NEWS	Houston	A B C	has been 147 years since this has been done, and Sam Houston probably needed fewer votes to win than we need signatures (to get on the ballot),	
3	1997	NEWS	WashPost	A B C	poor and transient and begin their education far behind their peers. Most years, fewer than half of the students in her school at the end of the school ye	
4	2007	MAG	ScienceNews	A B C	collected during weekends. Air samples collected during nighttime hours contained an estimated 13 percent fewer diesel emissions. Wood smoke was ar	
5	1991	MAG	Horticulture	A B C	would expect things to be better indoors. But they're not. There are fewer types of insect pests indoors, but they cause just as much damage. If	
6	2014	NEWS	Denver	A B C	. Among the factors: * The Colorado Parole Board granted an average of 24 fewer discretionary paroles per month, an 8 percent decrease. * Authorities	
7	2005	ACAD	NAJPsychology	A B C	empirical evidence. If a font is easier to read, then we predict that fewer attentional resources are spent attending to the process of reading and more at	
8	2011	NEWS	WashPost	A B C	of boys falling in love with banshees. # The conventional wisdom is that there are fewer books for boys because boys read less, because there are fewer	
9	1995	MAG	Ms	A B C	of arts funding. As money dries up, there will be more, not fewer , low-budget, intimate performances. Performance may also be used increasingly in corr	
10	2004	MAG	Money	A B C	just north of busy Clearwater, Dunedin shares the same beautiful waterfront, with far fewer tourist traps, plus a Main Street reminiscent of Mayberry. Hc	
11	2012				are unionized and their fragmented political voice increasing	
12	1990				rs. # A bill banning all teak imports from Thailand and Burma	
13	1995				er individuals were served in the 1991-1992 fiscal year. The nr	
14	2009				fewer than 500 strikeouts—in which case the original grouping	
15	2004	MAG	Redbook	A B C	. Instead, he said we should do a fine needle aspiration, which takes fewer cells, so that I'd have a basic pathology report to give to the	

Click here to see the paragraph, which gives you even more context to your search string.

Fig. 6: Concordances of fewer.

- Now repeat the steps but search for the word *less*. Compare its usage to that of *fewer* and try to formulate “rules” for the use of these words on the basis of their respective concordance lines.

Solution: ‘Less’ is used with uncountable nouns and with adjectives or adverbs, while ‘fewer’ is used with countable nouns.

- Additionally, you can ask students to pay special attention to register differences. For this, you might want to go back to step 5 and include more registers to allow a better comparison. Especially the inclusion of spoken and spoken-like text types (e.g. TV scripts or web language) could be interesting as this allows students to explore potential differences between conversational spoken and formal written English: are there any instances of *less/fewer* that deviate from the usage “rules” above? If yes, which nouns are most often affected by this? In which contexts and text types is this most common? See “Options and further ideas” below for a detailed explanation of how to compare the distribution of a particular construction in different registers using the “Chart”-function of the web interface.

Tasks

Pre-task

If your students are not yet familiar with using the concordance function of english-corpora.org, it is advisable to explain the procedure while working through the *fewer/less* example in plenum. Alternatively, you can choose any warm-up activity suitable for a lesson that asks students to remedy errors in their writing assignments.

Main task

In the main task of the lesson, ask your students to fill in the worksheet (see appendix) by querying the COCA. The worksheet consists of the following instructions:

- Write down three language issues that came up in your writing assignment.
- For each issue, find a sentence in the corpus which uses the word/phrase/construction in a suitable way and write it down.
- Describe how each word/phrase is used.

4. Use the information that you gathered from the corpus to rewrite your sentences.

Review questions

After having completed the main task, ask your students to answer the following review questions (cf. Friginal 2018: 229):

- Was using the concordance function helpful for learning more about word/phrase/grammar usage in English?
- In your opinion, is it worth spending the time learning how to use the concordance function?
- Are you likely to use the concordance function on your own in the future? Why/why not?

4 Options and further ideas

Variations

- Your students can also use their mobile phones instead of computers.
- As all of the research interfaces on *english-corpora.org* are similar, your students can apply their knowledge to explore other corpora, such as the British National Corpus (BNC), for instance.

Additional ideas

This section suggests further ideas on how your students can use corpus data to further improve their writing skills. If your students enjoy working with the COCA, it is highly recommendable for you to work through the following functions and present them one by one to your class. Should some of your students finish the tasks of the designed lesson earlier than expected, you may also use these ideas as gap fillers.

Exploring collocations

The term 'collocation' is used to describe words that tend to appear together. This is the case with *to make mistakes*, for instance. Thus, *mistake* is most frequently associated with, or collocates with, the verb *make*. By contrast, *to do mistakes* is considered unidiomatic because *mistake* does not collocate with the verb *do*. Exploring typical collocations involving different word classes, such as adjective + noun, or verb + noun, is another way of using the COCA in writing classes. Using the collocation function, you can see which words frequently co-occur with other words, which provides great insights into collocation usage and meaning. Here is an overview of how to use this function:

1. Open the COCA on www.english-corpora.org/coca, click on the '+'-icon and choose "Collocates" (Fig. 7).

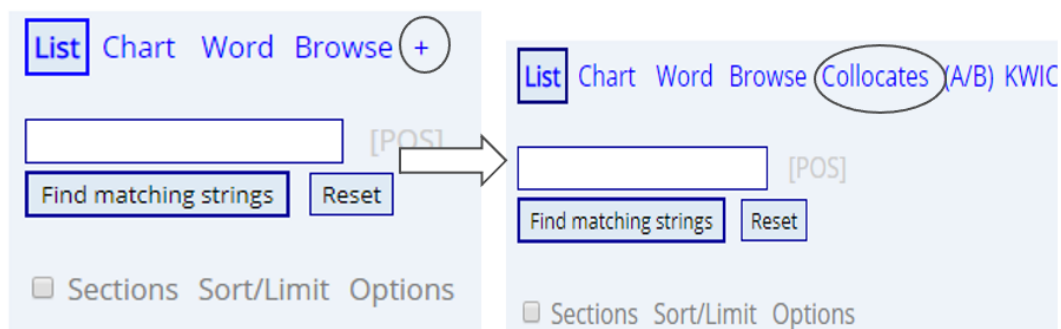


Fig. 7: Using the “Collocates” function.

2. You should now be able to view the following display (Fig.8):

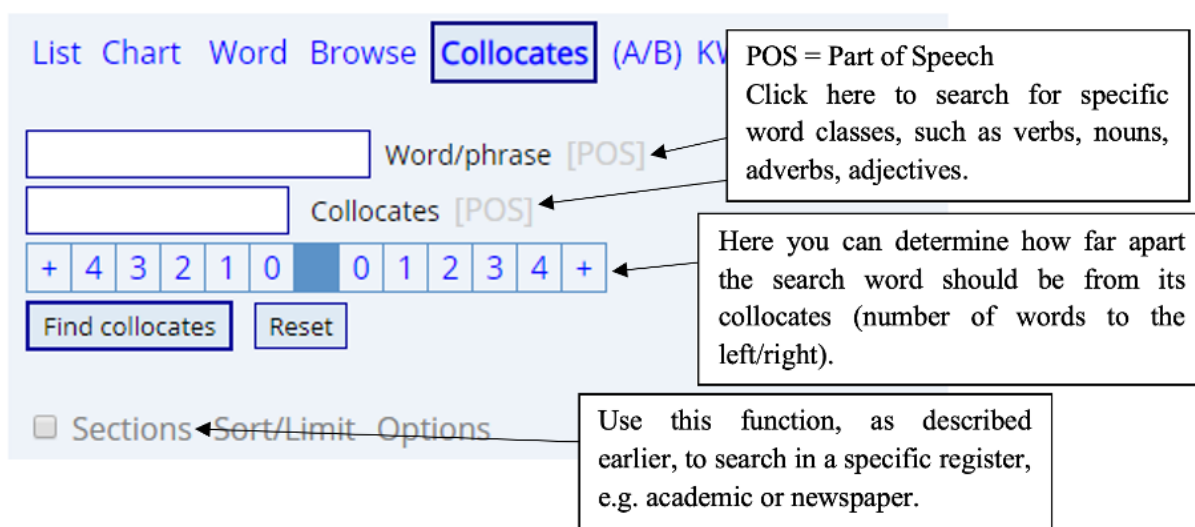


Fig. 8: Using the “Collocates” function.

Imagine the following scenario: you want to write a text about finances and are unsure about verbs that are commonly used with the noun *money*. Figure 9 shows an exemplary search for typical verbs one or two positions before the noun *money* in newspaper writing. Make sure to set the search options to “Group by lemmas” (not words) so that inflected forms such as *making*, *made* or *makes* will be classified as belonging to the same verb, in this case *make*.

The screenshot shows the COCA Collocates interface. The search term is 'money' and the part of speech is 'VERB'. The source list is set to 'NEWSPAPER'. The group by option is 'LEMMAS'. The results are sorted by frequency, with the top five verbs being [MAKE], [RAISE], [SPEND], [SAVE], and [GET].

1	<input type="checkbox"/>	[MAKE]
2	<input type="checkbox"/>	[RAISE]
3	<input type="checkbox"/>	[SPEND]
4	<input type="checkbox"/>	[SAVE]
5	<input type="checkbox"/>	[GET]

To see the context in which each verb occurs, simply click on it.

Fig. 9: Searching for typical verb collocates of the word money.

Thus, you can conclude that in newspaper writing, the five most frequently used verbs in front of *money* are *make*, *raise*, *spend*, *save* and *get*.

Searching for synonyms

Searching for synonyms is another possible way of using the COCA to increase the variety of word choice when writing texts. Open the COCA on www.english-corpora.org and choose "List". Type in an equal symbol (=), followed by any search string of interest. Figure 10 shows an exemplary search for synonyms of the adjective *beautiful*:

You can click on the [S] following any of these words to see the synonyms of that word.

You may also choose [SECTIONS] = [SHOW] in the search form to see the frequency of these synonyms in different genres, e.g. in newspaper or academic writing only.

The screenshot shows the COCA search interface. At the top, there are navigation tabs: List, Chart, Word, Browse, Collocates (A/B), KWIC. The search field contains "=beautiful" and a "[POS]" label. Below the search field are two buttons: "Find matching strings" and "Reset". Below these buttons are options: "Sections", "Texts/Virtual", "Sort/Limit", and "Options". To the right of the search interface is a list of 10 synonyms, each with a checkbox and a link to its synonyms page:

1	<input type="checkbox"/>	BEAUTIFUL [S]
2	<input type="checkbox"/>	WONDERFUL [S]
3	<input type="checkbox"/>	LOVELY [S]
4	<input type="checkbox"/>	ATTRACTIVE [S]
5	<input type="checkbox"/>	STRIKING [S]
6	<input type="checkbox"/>	HANDSOME [S]
7	<input type="checkbox"/>	GORGEOUS [S]
8	<input type="checkbox"/>	STUNNING [S]
9	<input type="checkbox"/>	CHARMING [S]
10	<input type="checkbox"/>	MAGNIFICENT [S]

Fig. 10: Synonyms of beautiful.

Exploring different registers

Searching in the COCA provides students with information about the frequency of words or phrases in different styles of English such as spoken, fiction or academic, for instance. Imagine the following situation: you need to write an assignment and you would like to know if it is appropriate to use the expression *freak out* in formal contexts. To answer this question, you have to follow three easy steps:

1. Open the COCA on english-corpora.org and choose the “Chart”-option, which shows the frequency of a search string in each section (Fig.11):

The screenshot shows the COCA search interface with the "Chart" tab selected. The search field is empty and has a "[POS]" label. Below the search field are two buttons: "See frequency by section" and "Reset". Below these buttons are options: "Sections", "Texts/Virtual", "Sort/Limit", and "Options".

Fig. 11: Using the “Chart” function.

2. Type *freak out* in the search field and press “See frequency by section”. You should get the following results (Fig. 12):

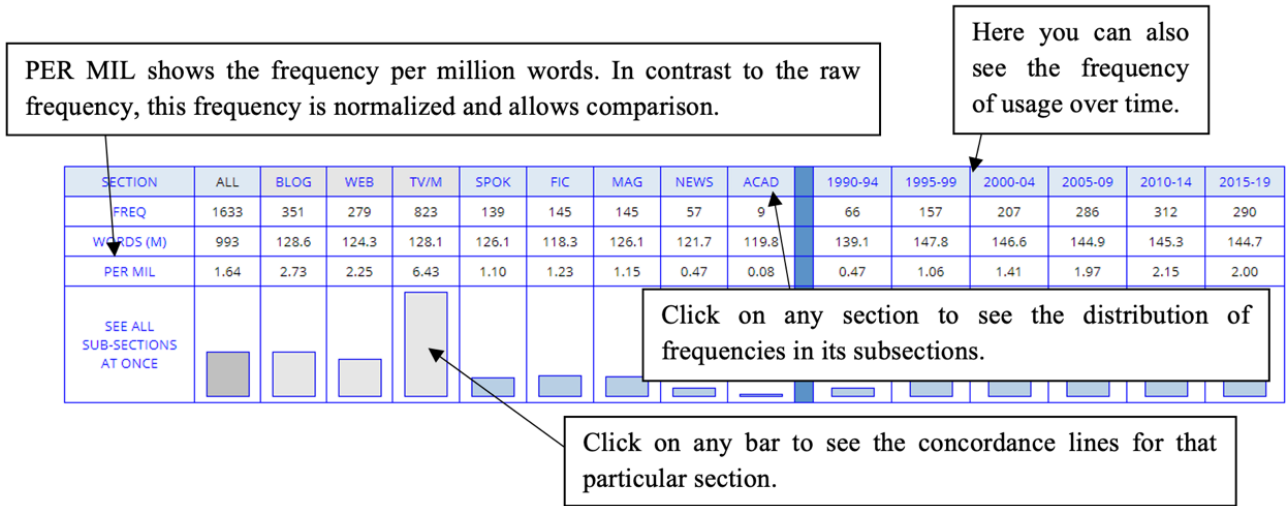


Fig. 12: The frequency of *freak out* in different registers.

3. Finally, you need to interpret the chart: Fig. 12 shows that the expression *freak out* in academic texts is only used 0.08 times per million words, which is the lowest value of all displayed sections. Thus, it is advisable to use a different expression in the assignment, such as the synonymously used term *to lose control*, for instance (see above: “Searching for Synonyms”).

The latest option of the COCA

In March 2020, the COCA was expanded in scope, size and features to make it even more useful for researchers, teachers, and learners, who can now browse through a list of the most frequent 60,000 words in the corpus, and see a wide range of useful information on each of these words (cf. Davies 2020b: 1). To make use of this, you can either click on “Word” and search for a specific word or click on “Browse” (see Fig. 13) to search by word form (e.g. *ism), rank order (e.g. words among the 6,000th, 26,000th or 46,000th most frequent words), pronunciation (e.g. words rhyming with *light*, or three-syllable words accented on the last syllable), or any combination of these (cf. Davies 2020b: 11).

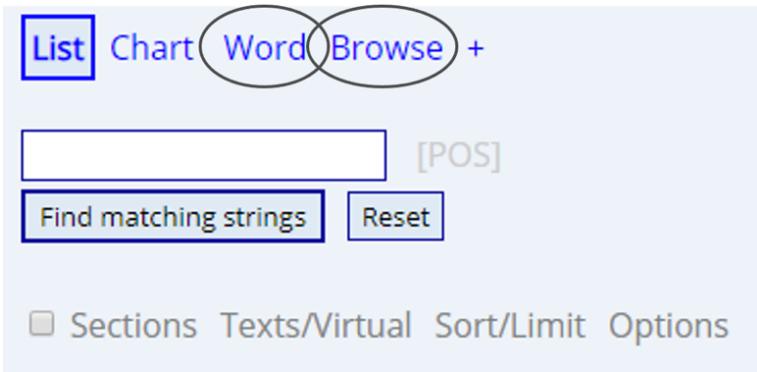


Fig. 13: Using the “Browse” function.

To explore this new feature, click on “Browse” and “Find random words” (Fig. 14).

HELP Find random words

Click on any of the search types below for more information and examples. (more...)

Word form	<input type="text"/>
Part of speech	<input checked="" type="checkbox"/> NOUN <input checked="" type="checkbox"/> VERB <input checked="" type="checkbox"/> ADJ <input checked="" type="checkbox"/> ADV <input checked="" type="checkbox"/> OTHER <input type="checkbox"/> ALL
Range	<input type="text"/> - <input type="text"/>
Pronunciation	Rhymes with <input type="text"/> Type <input type="text" value="EXACT"/>
Syllables / stress	○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ×

◀ ▶

Fig. 14: Using the “Browse” function.

	RANK	FREQ	Word	PoS	Audio	Video	Image	LANG?
1	19456	1044	tux	NOUN				
2	19457	1044	intoxicating	ADJ				
3	19458	1044	jasper	NOUN				
4	19459	1043	two-year-old	ADJ				
5	19460	1043	oppositional	ADJ				
6	19461	1043	moisten	VERB				
7	19462	1043	monotonous	ADJ				
8	19463	1043	frisbee	NOUN				

Fig. 15: Using the “Browse” function.

For each word on the list, you can hear the word pronounced, watch videos with the word in the text, see related images from Google Images, and get a translation for your preferred target language (Fig. 15). When you see a word of interest, click on it. You will be redirected to its homepage, which contains frequency information, definitions, translations, links to audio, images, and videos, as well as synonyms, topics (words that co-occur anywhere in the ca. 500,000 texts), collocates, clusters, and concordance lines (Fig. 18) (cf. Davies 2020b: 6). On the top right corner of each word’s homepage, you will find direct links to the respective pages of some of these features (cf. Davies 2020b: 8). For a more detailed account of the word’s collocates, clusters, concordance lines (KWIC) etc., simply click on the feature in question (Fig. 16). Figure 17 shows the collocation page of the word *trail*.

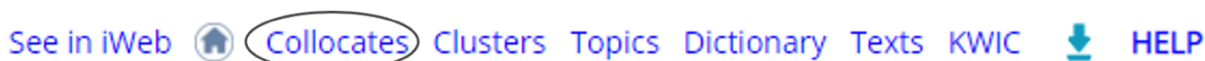


Fig. 16: Navigation tabs at the top of the word’s homepage.

+ NOUN	NEW WORD	?	+ ADJ	NEW WORD	?	+ VERB	NEW WORD	?	+ ADV	NEW WORD	?
3018	5.67	campaign	504	9.40	appalachian	1552	3.69	follow	18	4.27	left
1151	4.57	mile	264	5.82	steep	947	7.76	hike	17	3.69	downhill
851	3.65	paper	190	3.95	narrow	915	3.06	lead	14	5.32	steeply
623	8.61	blazer	159	3.95	historic	508	8.22	blaze	11	4.14	uphill
600	3.50	park	153	6.81	scenic	278	3.31	ride	7	3.17	aptly
514	3.89	mountain	132	5.47	rocky	178	3.32	climb	5	3.21	underfoot
470	2.94	blood	114	2.71	indian	173	2.75	cross	4	2.64	briskly
419	2.60	road	100	2.97	rough	154	3.78	wind	4	2.69	deftly
411	4.70	bike	86	4.51	mathematical	150	3.07	mark	3	2.57	northward
336	12.67	portland	82	7.02	paved	109	6.63	bike	3	2.64	eastward

Fig. 17: Collocates of trail.

trail (NOUN) See: VERB Rank #1-60,000 72192

Distribution across genres

Definition
1. a track or mark left by something that has passed 2. a path or track roughly blazed through wild or hilly country 3. evidence pointing to a possible solution

COLLOCATES (more) Words that co-occur nearby
NOUN campaign, mile, paper, blazer, park, mountain, blood, road
VERB follow, hike, lead, blaze, ride, climb, cross, wind
ADJ appalachian, steep, narrow, historic, scenic, rocky, indian, rough
ADV left, downhill, steeply, uphill, aptly, underfoot, deftly, briskly

SYNONYMS (more)
[path] road, track, path, trail, route, trajectory, footpath [track] trail, trace

CLUSTERS (more) 2, 3, 4 word strings
trail in - trail to - trail blazers - trail for - trail with - trails in - trail from - trail on
- trail campaign trail - paper trail - appalachian trail - hiking trails - long trail - blood trail - on trails - money trail
trail ... trail of tears - trail of blood - trail of destruction - trail that led - trail leads to - trail of smoke - trail of bread - trail of broken
- trail on the trail - along the trail - up the trail - off the trail - down the trail - leaving a trail - left a trail - to the trail
trail ... trail of bread crumbs - trail that led to - trail to the top - trail through the woods - trails in the sky - trail of destruction across - trail that leads to - trail of broken hearts
... trail on the campaign trail - out on the trail - from the campaign trail - end of the trail - hot on the trail - side of the trail - on the appalachian trail - miles of hiking trails

TEXTS / VIRTUAL CORPORA (more) Texts that have this as a "keyword"

WEB:nps.gov • WEB:ponderosaca.com • NEWS:AssocPress • MAG:Backpacker • ACAD:InstrPsych • MAG:Bicycling • NEWS:USA TODAY • MAG:Backpacker • FIC:FantasySciFi • MAG:Backpacker • MAG:NatIParks • NEWS:USAToday • MAG:Backpacker • MAG:Backpacker • WEB:...documentaryfilms.com • MAG:NatIParks • MAG:Bicycling • MAG:Backpacker • MAG:Bicycling • MAG:Backpacker • NEWS:Atlanta • MAG:Backpacker • MAG:Backpacker • MAG:NatIParks • NEWS:Atlanta •

CONCORDANCE LINES (more) The word in context (to see patterns of use)

1	WEB: 2012: amazon.com	Along the way , they confront Army Rangers along the Appalachian Trail and conduct late-night stealth mission to an abandoned NASA
2	BLOG: 2012: _heather.blogspot...	me , a direct definition of ultra trail running . The trails are amazing ! DCNR does an excellent job keeping the trails and
3	MAG: 1991: NatIParks	to protect a continuous corridor are ongoing . The North Country Trail Association is a 500-member organization established in 1980 ,
4	MAG: 1991: NatIParks	. # For more information , contact the Overmountain Victory Trail Association , c/o P.O. Box 632 , Manassas Park , VA 22111
5	MAG: 2003: Bicycling	and anti-bike website operator John Parulis , describing the trails at China Camp , east of San Rafael " Look at this
6	WEB: 2012: yelp.com	try ... # Perks of coming here is hiking the various trails at the top , and taking the ski lift back down .
7	BLOG: 2012: seattlemet.com	that the city did not " re-interview " businesses along the trail before approving its completion ; and the fact that an
8	TV: 2016: Tomorrow, When the W...	OK ? Rain gear 's there too . Should be some trail bikes out the back if they have n't been taken . Awesome
9	NEWS: 1998: Houston	because only a few of the wealthiest teams - the Portland Trail Blazers - owned by billionaire Paul Allen) , the New York
10	MAG: 2010: PopMech	, I also fail to notice that the side of the trail butts in the cliff off the creek bed . With my blissful

Fig. 18: Homepage of the word trail (Davies 2020b: 7).

It is hoped that this short introduction into the new "Browse"-option of the COCA gives valuable insights into its great potential for classroom use. Looking at the summary statistics page for each word, you can see a great wealth of information at once (collocates, clusters, synonyms, concordance lines and the distribution across genres), which saves a lot of time and effort and thus makes this new function very convenient. However, note that words that are not among the 60,000 most frequent words in the corpus do not have such a summary statistics page. Thus, you may well need to know about the original, basic functions of the interface to carry out your search. Starting with the basic functions gives students a better understanding of how corpora work and also gives them the necessary tools to search in all the other corpora featured on

English-corpora.org because the COCA is currently the only corpus on the web interface that supports this new function.

5 Caveats and limitations

After approximately 10-15 queries the website will ask you to register. Registration, however, is free to everyone and takes one or two minutes only. Afterwards you can carry on using the website as normal.



Fig. 19: Registering on the website.

Click on the yellow icon in the top right corner of the website (Fig. 19) to register. You should then see the following display (Fig. 20, Fig. 21):

1.

Fig. 20: Registering on the website.

2.

Name	<input type="text"/> <input type="text"/> (e.g. Mary Smith)	1. Type in your name.
Email address	<input type="text"/>	2. Type in a valid email address.
Password	<input type="password"/>	3. Choose a password.
Country	-- SELECT--	4. Select your country.
Category	<input type="radio"/> University professor: languages / linguistics <input type="radio"/> Graduate student: languages or linguistics <input type="radio"/> University professor: not languages / linguistics <input type="radio"/> Teacher: not university; not graduate student <input type="radio"/> Graduate student: not languages or linguistics <input type="radio"/> Student (undergraduate) <input checked="" type="radio"/> Other	3 RESEARCHER 2 SEMI-RESEARCHER 1 NOT RESEARCHER
	<input checked="" type="checkbox"/> I agree to the Terms and Conditions for this website	6. Please agree to the terms and conditions and press the "SUBMIT" button.
	<input type="button" value="SUBMIT"/> <input type="button" value="RESET"/> PROBLEMS ??	

Fig. 21: Registering on the website.

You will be asked to select your user status. Note that you must provide a web address that verifies your status in order to set up a Researcher or Semi-Researcher account. However, bear in mind that even if you are not approved as Researcher or Semi-Researcher, you can still use the corpora – just at a lower level of access (cf. Davies 2020a), which, however, is entirely sufficient for the purpose of this chapter. Within one or two minutes after having submitted your information, you will receive an email. Click on the link in that email and you will be able to continue using the web interface.

Occasionally, the following message (Fig. 22) will pop up, asking you to consider upgrading to a premium account. However, there is no need to upgrade your account. Please be patient and wait for a few seconds, after which you will be able to continue with your search.

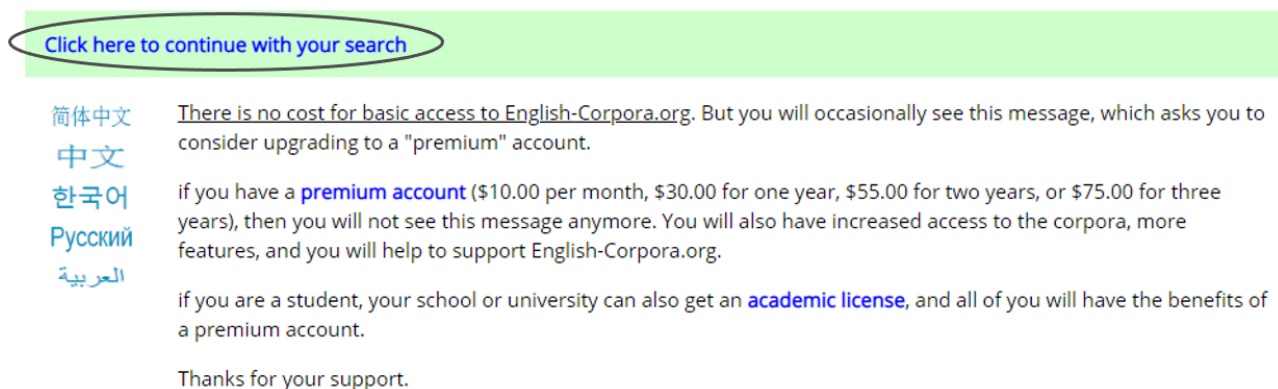


Fig. 22: Upgrade notification.

6 Conclusion

The main advantage of this lesson is that students are given the means to correct their mistakes independently with the aid of the COCA. Thus, students acquire learning strategies that enable them to solve problems in writing effectively by themselves, which meets one of the requirements of the curriculum. In addition, a recent meta-analysis of data-driven language learning interventions revealed that students are more likely to remember the language patterns they infer from their own corpus queries (cf. Boulton & Cobb 2017). Furthermore, the use of digital media in the classroom can contribute to higher motivation and engagement on the part of the students. On the teachers' side, there is not much preparation required for lesson, which is also beneficial. Moreover, the additional ideas mentioned above present many more interesting features of the web interface, which provides inspiration for many more COCA-based lessons.

7 Resources and references

Boerger, Claudia. 2020. Examples warm-ups. http://www.jochenenglish.de/misc/warm_up_beispiele.pdf (23 March, 2020).

Boulton, Alex & Tom Cobb. 2017. Corpus use in language learning: A meta-analysis. *Language Learning* 67(2). 348-393. <https://doi.org/10.1111/lang.12224>.

Davies, Mark. 2020a. English Corpora. <https://www.english-corpora.org/>

Davies, Mark. 2020b. The COCA corpus. https://www.english-corpora.org/coca/help/coca2020_overview.pdf.

Davies, Mark. 2008-2019. The Corpus of Contemporary American English (COCA): 600 million words, 1990-present. <https://www.english-corpora.org/coca/>

Friginal, Eric. 2018. *Corpus linguistics for English teachers: New tools, online resources, and classroom activities*. New York; London: Routledge.

Niedersächsisches Kultusministerium. 2015. Kerncurriculum für das Gymnasium, Schuljahrgänge 5-10: Englisch. <https://cuvo.nibis.de/cuvo.php?p=download&upload=139> (12 October, 2020).

Wichmann, Anne, Steven Fligelstone, Tony MacEney & Gerry Knowles (eds.). 1997. *Teaching and language corpora*. London: Pearson Education.

8 Appendix

Worksheet: How can I improve my writing?

Name: _____ Date: _____

Write down three issues that came up in your writing.	For each issue, find two sentences in the corpus, which use the word(s)/phrase(s) you struggled with and write them down.	Describe how each word/phrase is used. Copy the example sentences that you think are most likely to help you.	Use the information that you gathered from the corpus to rewrite your sentences.
1			
2			
3			

This worksheet was downloaded from: Busmann, Katharina (2021). "How to do data-driven learning (DDL) with secondary school pupils to improve their writing." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

Example Worksheet: How can I improve my writing?

Name: _____ Date: _____

Write down three major issues that came up in your writing.	For each issue, find two sentences in the corpus, which use the word(s)/phrase(s) you struggled with and write them down.	Describe how each word/phrase is used. Copy the example sentences that you think are most likely to help you.	Use the information that you gathered from the corpus to rewrite your sentences.
<p>1 Confusion of <i>since</i> and <i>for</i></p> <p>I know her <i>since</i> a long time.</p> <p>She has been talking on the phone <i>for</i> dinner.</p>	<p>I have been in the recruitment game <i>for a long time</i>.</p> <p>He had not bothered to eat or drink <i>since</i> dinner with Georgia the night before.</p>	<p><i>For</i> refers to a length of time/period of time. <i>Since</i> refers to a specific point in time.</p> <p>Both are most often used with a verb in the present perfect or present perfect progressive.</p>	<p>I've known her <i>for</i> a long time.</p> <p>She has been talking on the phone <i>since</i> dinner.</p>
<p>2 Idiomatic preposition use</p> <p>At Monday she went to the doctor.</p> <p>I visited her <i>at</i> Wednesday.</p>	<p><i>On</i> Monday, he paid the back taxes.</p> <p>The fight occurred <i>on</i> Wednesday.</p>	<p>The preposition used with days of the week is <i>on</i>. If the sentence starts with the weekday, there is often (though not always) a comma afterwards.</p>	<p><i>On</i> Monday, she went to the doctor.</p> <p>I visited her <i>on</i> Wednesday.</p>
<p>3 Used to + gerund / infinitive</p> <p>Nowadays people don't have enough time to cook so they <i>get used to</i> buy pizza or hamburgers.</p> <p>Music <i>used to</i> playing a big role when I was fourteen.</p>	<p>Neighbors and acquaintances say the man and the woman <i>used to quarrel</i>.</p> <p>My mother is <i>used to getting</i> what she wants.</p>	<p><i>Used to + infinitive</i> refers to something that happened regularly or repeatedly in the past but now doesn't happen anymore.</p> <p><i>To be/get used + -ing</i> refers to habits or thought patterns, which are ongoing.</p>	<p>Nowadays people don't have enough time to cook so they <i>get used to buying</i> pizza or hamburgers.</p> <p>Music <i>used to play</i> a big role when I was fourteen.</p>

This worksheet was downloaded from: Busmann, Katharina (2021). "How to do data-driven learning (DDL) with secondary school pupils to improve their writing." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

7. Teaching neologisms using corpus tools

NICOL SPERLING AND SARAH WINKELMANN

1 Introduction and rationale

This chapter combines teaching stylistic devices, specifically neologisms, with corpus linguistics in the EFL classroom. It proposes a lesson plan in which methods and topic choice are strongly oriented towards Lower Saxony's Core Curriculum for Secondary Education, which emphasizes the importance of acquiring meaning through "intelligent guessing" and contextualization (Niedersächsisches Kultusministerium 2015: 16). It also states that students should be able to analyze and interpret literary and non-literary texts, e.g., by making use of their knowledge of stylistic devices (Niedersächsisches Kultusministerium 2017: 21). Stylistic devices are more frequently covered in the upper classes but can occasionally be discussed in lower grades, too. This lesson plan can serve as an introduction to teaching stylistic devices while at the same time helping students to develop "intelligent guessing" strategies, creative writing and reading comprehension skills.

In addition to raising your students' awareness of neologisms and how language constantly evolves, this chapter encourages them to familiarize themselves and work autonomously with a corpus. It provides you with a lesson plan and a step-by-step guide to better understand and make use of the functions of a corpus in relation to the topic. It also includes suggestions on how to use additional resources, such as [Wordspy.com](https://www.wordspy.com), a website that provides you with novel words and phrases which have recently come into use in English.

2 Outline and objectives

This chapter prepares you for giving the proposed lesson by ...

- showing you how to use the corpus tool [english-corpora.org](https://www.english-corpora.org) and [Wordspy.com](https://www.wordspy.com) (~ 30 min).
- familiarizing you with the lesson plan (~ 10 min).
- providing you with suggestions on how to create corpus-based material for the proposed lesson (total of 1-2 hours; duration partly depends on how much the worksheet template needs adjustments to suit your learner group).

The proposed lesson is suitable for secondary school students aged 15 or older (10th grade according to the German school system)

- Lesson unit: Teaching stylistic devices
- Topic: Neologisms (90 minutes)
- Main competence: reading comprehension (contextualization and intelligent guessing)

Learning Objectives

After this lesson, students ...

- can search for words and collocates using english-corpora.org.
- can systematically navigate through text sample results from a corpus.
- can work independently with a corpus and use it as a glossary and source of information.
- can deduce the meaning of unknown vocabulary (neologisms) used in different contexts, through intelligent guessing.
- can write their own definitions of previously unknown words on the basis of their corpus findings.

3 Corpus, corpus tools and methods

The proposed lesson makes use of the corpus website english-corpora.org as an analysis tool for students and teachers. It is freely accessible, which makes it particularly attractive for classroom use. Additionally, the handling of a corpus is easy to learn, as various tutorials can be found on the internet. Some of the main features of this corpus tool are the possibility to generate word lists, to display frequencies as charts and to find frequent collocations.

This chapter uses the NOW (News on the Web) corpus, which can only be found on english-corpora.org. Since neologisms are constantly changing and developing, especially in a digitalized age, a corpus which is regularly expanded by the addition of new content from articles and blog posts can be a useful classroom resource. Monthly content updates (about 180-200 billion words) put the NOW corpus, as of March 2021, at about 12.2 billion words. It is the largest and most up-to-date corpus found on english-corpora.org and also very comprehensive and differentiated, as it allows users to set particular search criteria such as year of publication and country. The NOW corpus also provides a regularly updated list of words and phrases provided by WordSpy.com, a constantly updated website of the most recent neologisms.

The proposed lesson enables students to contextualize neologisms by employing intelligent guessing strategies and using their existing knowledge with the help of the collocate function on english-corpora.org.

4 Teacher preparation

4.1 Step-by-step guide on using the corpus tool

The NOW corpus can be accessed by registering on english-corpora.org.

The most widely used online corpora: [guided tour](#), [overview](#), [search types](#), [variation](#), [virtual corpora](#)

The links below are for the online interface. But you can also download the corpora for use offline.

Corpus (online access)	Download	# words	Days	Genre(s)
iWeb: The Intelligent Web-based Corpus		14 billion	6 countries	Web
News on the Web (NOW)		12.6 billion+	20 countries	Web: News
Global Web-Based English (GloWbE)		1.9 billion	20 countries	Web (incl blogs)

Account menu: Register / profile, Log in, Log out, Name of university, Reset password, Delete account

Fig. 1: Corpus overview (english-corpora.org)

You can create a free account under “Register”, which can be found if you go to “My account”. This is also where you log in afterwards (Fig. 1).

If you click on “Corpora” after logging in, you will be provided with a shortcut to an overview of all the corpora listed on english-corpora.org. You can now choose the NOW corpus (Fig. 1).

Search interface: List, Chart, Collocates, Compare, KWIC. [POS]? Find matching strings, Reset.

Search by date: Sections Texts/Virtual Sort/Limit Options

Logged in status: **LOGGED IN**

See new features for the KWIC/concordance page

Download the corpus for offline use

The NOW corpus (**N**ews on the **W**eb) contains **12.2 billion words of data** from web-based newspapers and magazines from 2010 to the present time (the most recent day is **2021-03-25**). More importantly, the corpus grows by about **180-200 million words of data each month** (from about 300,000 new articles), or about two billion words each year.

While other resources like [Google Trends](#) show you what people are *searching for*, the NOW Corpus is the only structured corpus that shows you what is actually *happening* in the language – virtually right up to the **present time**. For example, see the **frequency of words** since 2010, as well as **new words and phrases** from the last few years.

Click on any of the links in the search form to the left (such as [List](#) or [Chart](#)) for context-sensitive help, and to see the range of queries that the corpus offers. You might pay special attention to the **comparisons** between dates and countries and **virtual corpora**, which allow you to create personalized collections of texts based on (sub-)register, website, and even words in the web pages.

[Five minute tour](#)

Fig. 2: NOW corpus overview (english-corpora.org/now/)

Once you have clicked on the NOW corpus, the interface provides you with different features, as shown in Fig. 2. The left-hand side will be discussed later, when describing the tasks designed for the students. The right-hand side provides you with information on whether you are logged in on english-corpora.org. Further down, you will see a shortcut to “New words and phrases”. It will take you to a table of words and phrases which have recently been added to the English language and can therefore be used as a guide when choosing neologisms for your individual worksheet.

These are some words and phrases that have recently come into English. The list comes from WordSpy.com, where you can find more than 3,000 other words as well.

Culture / sociology / The World	
Brexit (n)	The exit of Great Britain from the European Union.
wokeness (n)	The state of being aware of and sensitive to social justice issues.
feebate (n)	A government program designed to reduce energy use and pollution by levying a fee on fuel-inefficient vehicles and offering a rebate on fuel-efficient vehicles.

Fig. 3: Examples of new words and phrases (english-corpora.org/now/)

The table is divided into two different categories: “Culture/sociology/the World” and “Computers/technology/science” (Fig. 3). It contains the word and a brief definition. By clicking on the word, you obtain information on its frequency of use by year and country.

Since this list is provided by WordSpy.com, you will also find a link to this website. Here is a brief overview of what it looks like and instructions on how to find words and phrases.

WORD SPY

By Alpha

By Tag

By Date

Random Word

Top 100

Full Search

Word search

WORD SPY

THE WORD LOVER'S GUIDE TO NEW WORDS

Most Popular Words

1. **biostitute**
2. **skinship**
3. **work-life balance**
4. **walk the cat back**
5. **kittenfishing**
6. **throuple**

Fig. 4: Main page of Wordspy.com

WordSpy.com lists over 3,000 words and phrases including a brief definition, related words, etymology and examples of its usage. The website itself is well structured. The menu on the left side (Fig. 4) offers different ways to search for words, e.g., the option “By Date”, which arranges words and phrases by year and month. Another way to find new words is the option “Random Word”, which will show you a new neologism or phrase each time you click on it. It is also possible to conduct a more detailed search by selecting “Full Search” or to search in different categories (“By Tag”).

glamping

n. A form of camping that includes expensive equipment, fine food, and other luxuries.

Other Forms

glamper *n.*

Etymology

glamor+ *camping*

Fig. 5: Details of the neologism glamping

2005 (earliest)

These days it's more "**glamping**" than camping, with the best companies offering state-of-the-art pre-erected tents and luxe mobile homes with ensuite bathrooms that feel like an Oscar-winner's trailer.

—Susan Ward Davies, "Know before you go," *The Guardian*, February 19, 2005



Fig. 6: Example sentence glamping

Once you have chosen a word you can simply click on it and Wordspy.com will present you with a more detailed overview of morphologically related words, an account of the word's etymology (Fig. 5), and various examples of how it is used (Fig. 6).

When selecting a word, it is crucial to check whether it occurs with sufficient frequency in the corpus for students to work with.

List Chart **Collocates** Compare KWIC

glamping Word/phrase [POS]?

NOUN Collocates noun.ALL  

+ 4 3 2 1 0 0 1 2 3 4 +

Find collocates Reset

Search by date

Sections Texts/Virtual Sort/Limit **Options**

HITS

KWIC

GROUP BY

DISPLAY

SAVE LISTS

Fig. 7: Search settings in the collocates function

When you go back to english-corpora.org, once again choose the NOW corpus and click on “Collocates” (Fig. 7). First, type in the neologism you want to analyze, in this case *glamping*, and select the parts of speech you are interested in. This can be done by clicking on “Collocates [POS]”. If you only want to find examples of the word being used as a noun, choose the corresponding category from the dropdown menu (Fig. 7). To search for all word classes, type in a star (*). Second, select the “span”, i.e., the distance between the search word and its collocates; if you want to know which words typically occur immediately before or after the word, for instance, select position “1” on both sides (Fig. 7). Third, change the search settings to “Group by Lemmas” and “Display Per Mil”. You can now click on “Find collocates” to get results.

SEARCH			FREQUENCY		CONTEXT		CONTEXT +
ON CLICK: CONTEXT TRANSLATE (??) GOOGLE IMAGE PRON/VIDEO BOOK (HELP)							
HELP			FREQ	PER MIL			
1	<input type="checkbox"/>	[SITE]	283	0.00			
2	<input type="checkbox"/>	[EXPERIENCE]	215	0.00			
3	<input type="checkbox"/>	[TENT]	206	0.00			
4	<input type="checkbox"/>	[POD]	139	0.00			
5	<input type="checkbox"/>	[OPTION]	93	0.00			
6	<input type="checkbox"/>	[ACCOMMODATION]	61	0.00			
7	<input type="checkbox"/>	[TRIP]	52	0.00			
8	<input type="checkbox"/>	[RESORT]	51	0.00			
9	<input type="checkbox"/>	[VILLAGE]	44	0.00			
10	<input type="checkbox"/>	[HUB]	40	0.00			
11	<input type="checkbox"/>	[LUXURY]	40	0.00			
12	<input type="checkbox"/>	[BUSINESS]	38	0.00			
13	<input type="checkbox"/>	[SPOT]	37	0.00			
14	<input type="checkbox"/>	[RETREAT]	33	0.00			
15	<input type="checkbox"/>	[FACILITY]	31	0.00			

Fig. 8: Frequency list of the collocates of glamping

As a result, the corpus displays a list of collocates sorted by frequency (Fig. 8). Note that there is no distinction between the preceding and the following word. To find out whether a collocate typically occurs to the left or right of the search term, click on the word in question. This will provide you with more context.

SEARCH			FREQUENCY		CONTEXT		OVERVIEW
FIND SAMPLE: 100 200 PAGE: << < 1 / 3 > >>							
CLICK FOR MORE CONTEXT				EXPLORE NEW FEATURES		SAVE TRANSLATE ANALYZE	
1	21-03-24 GB	dorsetecho.co.uk				. Undercover Woodland Treehouse, West Sussex # Take to the trees with this elevated glamping experience . Housed in an ancient oak tree in West Su	
2	21-03-21 US	Yanko Design				camping experience once in a while, or to be more millennial, a comfortable glamping experience once in a while! But what about our feline friends? \	
3	21-03-18 SG	sg.asiatatler.com				their own versions of it. From picnics at home, with a view to glamping experiences , here's what's in store for you if you want to go	
4	21-03-17 GB	cityam.com				Spiral Straw Bale Studio, near Canterbury # The Buzzardry, East Sussex # Unique glamping experience , near Ely & Cambridge # Undercover Woodland	
5	21-03-17 GB	cityam.com				of all it's right in the heart of the South Downs. Unique glamping experience , near Ely & Cambridge # Finishing with a flourish, this one looks	
6	21-03-17 GB	independent.co.uk				West Sussex # Undercover Woodland Treehouse Airbnb # Take to the trees with this elevated glamping experience . Housed in an ancient oak tree in'	
7	21-03-17 SG	sg.asiatatler.com				stays in parks, cabins, yurts, treehouses and RVs, as well as glamping experiences . Currently only operating in the US, it has acquired an Australian co	
8	21-03-11 US	USA Today				the ground, either. # " So many campgrounds now offer cabins and other glamping experiences . " Rossi says, noting that if you've never tried RV trav	
10	21-02-23 US	CNN (1)				south of Sri Lanka near Yala National Park. Back of Beyond offers a glamping experience overlooking sprawling sand dunes and natural pools. # " Our	

Fig. 9: Context results for the collocate experience

Here, we chose the collocate *experience*. The website provides you with a list of example sentences from news websites, along with additional information on the source, country, and date (Fig. 9). To see more context, click on the corresponding number on the left-hand side, which redirects you to a larger excerpt of that article, including the full source and title (Fig. 10).

SEARCH	FREQUENCY	CONTEXT	CONTEXT +
--------	-----------	---------	-----------

Source information:

Source	https://www.yankodesign.com/2021/03/05/camping-tents-that-meet-all-your-modern-millennial-glamping-needs/
Country/date	US (21-03-21)
Title	Camping tents that meet all your modern millennial Glamping needs!

Expanded context:

same time let guests be immersed in it. The campground includes buildings featuring a kitchen, common gathering spaces, a fire pit, and more. All the facilities were carefully constructed by consciously choosing to make it environmentally friendly -- lightweight structures used sustainably sourced timber, an on-site sewer treatment to avoid adding waste to natural water bodies, collecting rainwater, having passive solar access, cross-ventilation to reduce power consumption, growing and preserving vegetation and more. # We all like a comfortable camping **experience** once in a while, or to be more millennial, a comfortable **glamping experience** once in a while! But what about our feline friends? Wouldn't they like a warm tent to cozy up into? I mean maybe they haven't ever thought of it... but I'm sure they would enjoy it! Well, the Cat Camp is here! And it's a mini indoor tent, especially for our cats. It's quite similar to the mini display tents seen in outdoor shops, and it's exactly like the larger tents for humans except its tiny... and

Fig. 10: Extended context for glamping experience

You can always choose to switch between “Search”, “Frequency”, “Context”, and “More context” on the navigation bar above.

4.2 Customizing the worksheet template

Below you will find a worksheet template and an answer sheet, created specifically for the proposed English lesson. This material was created in February 2021 and is targeted at German 10th grade students of the Gymnasium (a type of secondary school in Germany). The worksheet provides the students with three main tasks outlined in the lesson plan. At this point, it is important to consider the language proficiency of your learner group. The meanings of the neologisms on the worksheet might be too obvious or they might be outdated. Considering ongoing language change, especially the introduction and spread of new words, it is crucial to scrutinize the neologisms on the worksheet. Finding more suitable words should not be too difficult thanks to the very many neologisms listed on Wordspy.com and our guide on how to use the website.

Worksheet

- 1) Work with a partner and have a look at the neologisms listed in the table. Make some guesses about their meaning and note them down in the column "Guesses".
- 2) Together with your partner, go to the website english-corpora.org and select the NOW corpus. Use the step-by-step-guide "How to use english-corpora.org" to complete the upcoming tasks. Work through steps a-d for each neologism and document your results in the table.
 - a) Search for the neologism and scan the results. Pick an example sentence which includes the unknown word. Add the source of the quoted sentence.
 - b) Use the collocates function to find the word's most frequent collocates (the words that typically occur in combination with the word you searched for).
 - c) Select different collocates from the list and examine the sample sentences to figure the meaning of the neologism. Keep in mind that a word might have more than one meaning (key words are enough).
 - d) Based on your previous search results (example sentences, collocates, meaning/s), write a suitable definition of the neologism. Use your own words.
- 3) Find another group who has also finished the tasks above and compare your results. Did you make similar guesses before using the corpus tool? Are your definitions similar or totally different?

Answer Sheet

Neologism	Guesses	Example Sentence and Source	Collocations	Meanings	Definition
Glamping	Upcycling lamps	„The options range from modest, authentic tented camping to full-on luxury glamping .” (Winnipegfreepress.com)	Glamping sites Glamping experience Glamping tent	Glamorous camping	A luxurious form of camping which includes fine food and expensive equipment.
Superager	Elderly people Old car (oldtimer) Good and old wine	„A superager’s brain connectivity looks like a younger adult in fMRI.” (psychologytoday.com)	Superager’s cortex Superager’s brain	Super aging Aging while still having a well-functioning brain	Individuals over the age of 80 who have the same brain capacity as younger people.
Butterfingers	Cookies Buttery and doughy fingers Sticky fingers	“Here’s an ideal case of anyone with butterfingers. It’s made from a soft yet durable TPU material, which [...] also helps you grip it.” (digitaltrends.com)	With butterfingers Their butterfingers In case of butterfingers	Buttery fingers Being clumsy Letting things slip out of one’s hands easily	Someone with slippery and clumsy fingers, often failing to hold on to objects.
Furkid	Person stuck in the phase of a child/teenager Kid seen as a pet	“Even if your furkid has short fur, that only buys you some time between grooming sessions.” (asiaone.com)	Your furkid My furkid Beloved furkid	Furry kid Pet seen as a child/kid	A pet that means as much to its owner as an actual child.
Begpacker	Homeless Begging for money to travel	“Just yesterday, Malaysians were horrified at a disturbing video of a foreign ‘begpacker’ couple swinging their baby in the air to entertain locals [...] in hopes that they would receive donations for their travels.” (worldofbuzz.com)	The begpacker Begpacker problem Begpacker phenomenon	Begging backpacker	A backpacker begging for money to travel.

5 Lesson plan

5.1 Outline

The plan is designed as a 90-minute lesson for 10th grade students in secondary school. It can be implemented within or at the beginning of a teaching unit on analyzing literary works/stylistic devices. The preceding lesson may be used to introduce corpus tools and methods in order to save time during the planned lesson. You may also want to give an introductory lesson on the use and function of the most common stylistic devices prior to the proposed lesson.

We advise you to begin the lesson without telling the students what the topic will be. The teacher starts by showing them a chat from 2021 filled with neologisms, all linked to the Covid-19 pandemic (Fig. 11).

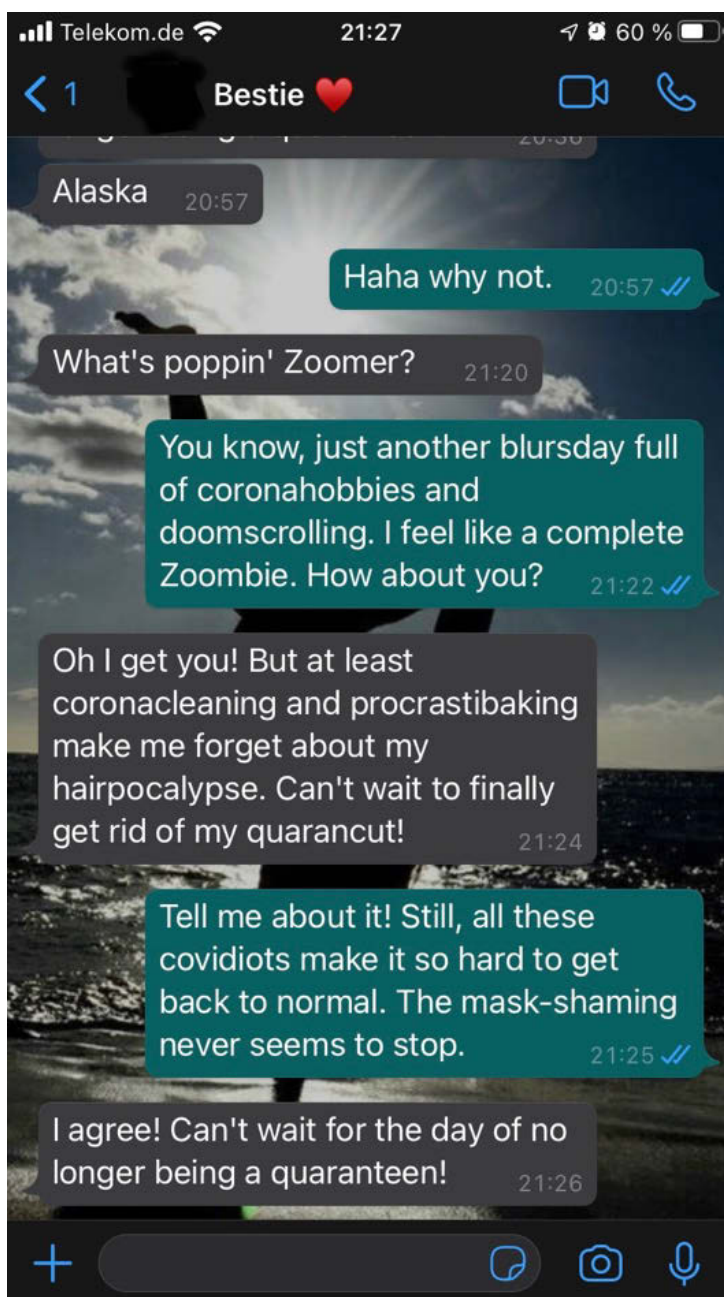


Fig. 11: WhatsApp chat 2021 (picture by the authors)

After two students have read the chat out loud, the teacher asks the class what the conversation is about and what is striking about the choice of words. Students may or may not be familiar with neologisms. It

is the teacher's task to use the neologisms of the chat to initiate a class discussion and to activate the students' prior knowledge, encouraging them to work out the definition and function of this stylistic device. The pandemic can be used as a concrete example of how language evolves and changes over time. As a result of new situations and events, novel words and phrases are created and become part of language use. However, these newly created words do not always remain neologisms. Over time, they either vanish or are added to the dictionary, as you can see by our example of *glamping*.

The main part of the lesson focuses on working with the NOW corpus on english-corpora.org. Students are provided with a worksheet which features a list of neologisms (see 4.2). The first task is to get together with a partner and to speculate about possible meanings of the listed words. Some of these suggestions can be collected on the blackboard in order to review them at the end of the lesson. One aim is to enable students to figure out the meaning of a word through contextualization and intelligent guessing or by analyzing its context. To further explain why such strategies are useful, you can point out that most neologisms will not be found in a dictionary. Up-to-date corpora, however, might include words which have not yet been taken up by dictionaries. Subsequently, the NOW corpus and the corpus website are introduced to the students. Here, it is important to consider whether your students have already worked with english-corpora.org or not. If not, make sure to give them a brief explanation and guide them through the different functions. With the help of a step-by-step guide for using the corpus tool (see Appendix), students work in pairs and complete the tasks on the worksheet. They search for the given neologisms in the NOW corpus, find texts which include these neologisms as well as their most frequent collocates, and discover their meaning by analyzing the context. In addition, they note down an example sentence for each neologism and formulate their own definitions of the previously unknown words.

The findings are discussed and recorded by, firstly, comparing the results in groups of four; secondly, by presenting them in class; and thirdly, by contrasting them with the assumptions made by the students at the beginning of the lesson. After the lesson ends, the final results should be made available to all of the students, which is why you should note them down during the class discussion.

Additional tasks and ways to adjust the lesson according to individual needs and preferences are presented in section 6.

5.2 Lesson plan

Procedure	Content	Arrangement	Media/ Material
Introduction ~ 10 min	<p><u>WhatsApp Chat from 2021</u></p> <ul style="list-style-type: none"> The chat is displayed for everyone to see and read out loud by two students T starts a class discussion about the chat, asking questions about its content and the neologisms used, e.g.: <ul style="list-style-type: none"> – “What age group could be represented in the chat?” – “Can you spot unfamiliar words?” – “What could they mean?” – “Can you identify the stylistic device that is used here?” Class discussion about the most significant neologisms found in the chat An answer sheet with explanations of all the neologisms used in the chat will be projected onto the wall and handed out to the students (see Appendix) Revision of the definition and function of neologisms as a stylistic device T asks if students can think of any other neologisms recently added to the English language 	Plenum	<p>Projector, laptop</p> <p>Answer sheet “Corona neologisms and their meaning”</p> <p>Blackboard</p>
Main part ~ 10 min	<p><u>Guessing the meaning of neologisms (task 1)</u></p> <ul style="list-style-type: none"> Students are asked to think of possible meanings of several neologisms A few proposals are collected Impulse question: “How could we find out what a word means if we cannot find it in the dictionary?” 	<p>Partner work</p> <p>Plenum</p>	<p>Worksheet</p> <p>Blackboard</p>
~ 15 min	<p><u>Introducing the corpus tool</u></p> <ul style="list-style-type: none"> Brief explanation of the NOW corpus and why it is beneficial to use corpus tools in this lesson T simulates a word search on english-corpora.org 	Plenum	Laptop/tablet, projector
~ 35 min	<p><u>Working with the corpus tool</u></p> <ul style="list-style-type: none"> T explains the tasks on the worksheet Students work with the corpus tool and search the NOW corpus to find texts which include the neologisms on the worksheet Students analyze the results and work out the meaning of the neologisms by having a closer look at the contexts and collocations of the words Students copy a sentence as an example and write a possible definition of each neologism based on their search results and interpretations 	<p>Plenum</p> <p>Partner work</p>	<p>Worksheet, step-by-step guide for students, laptops/ computers</p>

<p>Conclusion</p> <p>~ 20 min</p>	<p><u>Comparing results</u></p> <ul style="list-style-type: none"> • Two groups (total of 4 students) come together to compare their results • Presentation and discussion of results • Comparison with the guesses made at the beginning of the lesson • T fills out the worksheet with the collected results for the whole class to see 	<p>Groups of 4</p> <p>Plenum, message chain</p>	<p>Worksheet</p> <p>Document imaging camera/ laptop, projector, worksheet</p>
-----------------------------------	---	---	---

Table 1: Lesson plan; T= Teacher

6 Options and further ideas

Different proficiency levels require different activities and methods. You might want or need to adapt specific parts or activities of the lesson according to your individual learner group. This section proposes possible alternatives and additions as well as impulses for modifying the lesson to get the most out of it for your students and to suit your teaching style.

As it was already mentioned, it should be taken into consideration that an introduction to english-corpora.org and corpora in general prior to the lesson could make it easier for some students to follow the corpus activities, while also saving time during the main part of the lesson. However, we assumed that there is usually not enough time to integrate an introductory lesson into your teaching unit. Therefore, we included an introduction to the NOW corpus and an explanation of the corpus tools in the lesson plan. Nonetheless, we still advise you to reserve an earlier lesson to familiarize students with the basics.

Moreover, specific parts of the lesson plan could be replaced with alternatives. For instance, another option for the warm-up activity could be to provide the students with a list of neologisms and make them guess their meanings right away. This could also be done in the form of a challenge or game, e.g., by creating teams and collecting guesses on the blackboard. The students would then be required to use the corpus tools to investigate which guesses were right and which ones were wrong. Another option would be to initiate a class discussion about what neologisms are and to collect examples of recent English neologisms which the students might have heard of. This approach activates the students' existing knowledge on stylistic devices.

Additional tasks can be prepared for very advanced students. The corpus tool provides far more features which can be examined by the students. The chart function, for instance, allows you to see in which years a neologism appeared most frequently and also lets you search for the frequency of a word in different countries. This could be useful for finding out when a neologism came into being and how it is used in different varieties of English. The meaning or frequency of use of an English neologism may well differ between news texts from India and the United States. For instance, in India the neologism *butterfingers* occurs 0.08 times per million words, whereas in the United States it is found 0.03 times per million words (NOW Corpus). Students could investigate why this might be the case. When analyzing the Indian text samples, it becomes evident that the word is often used in reference to the comic book character "Butterfingers", created by Khyrunnisa A., who is a famous Indian author of children's fiction. By contrast, a closer look at the texts published in the United States reveals that "Butterfingers" is also the name of a famous candy bar. However, both countries also use the word to describe a "person who drops things they are carrying or trying to catch" (dictionary.cambridge.org). The chart function, which displays a word's

frequency sorted by years, can be used to analyze in what year a neologism first appeared and when it was most frequently used. For instance, the chart of the NOW corpus shows that the use of the word *Brexit* started to increase in 2016, when the results of the Brexit referendum were published and discussed all over the media (see Fig. 12).

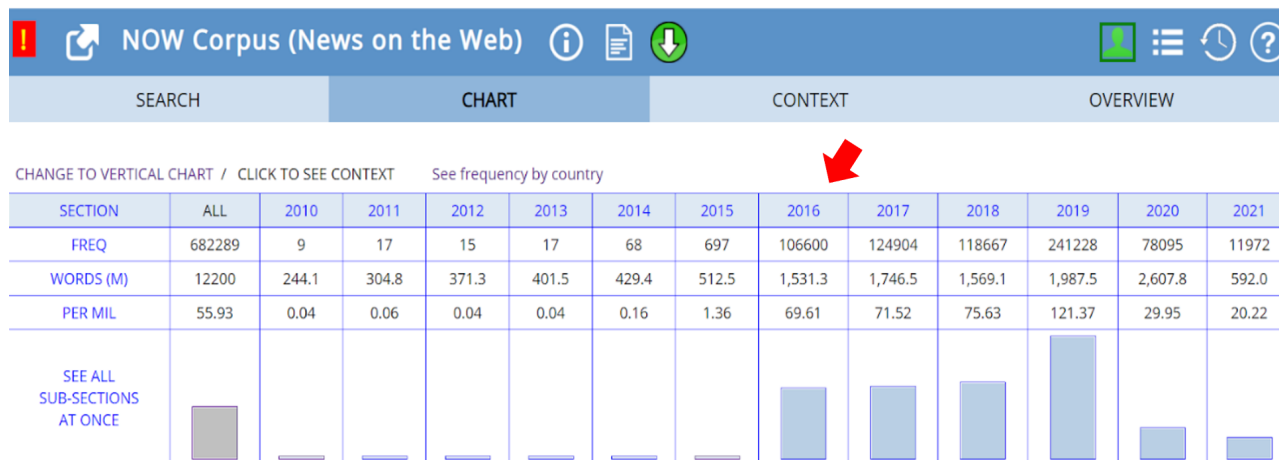


Fig. 12: Chart with frequency results for *Brexit*

These examples show the variations in meaning and usage of a word, as well as its connection to past historical and cultural events, which can be included into your lesson. Students could either work autonomously and follow their personal interests once they have finished the tasks on the worksheet, or you could choose which words/topics might be relevant to investigate further.

Lastly, we believe that the corpus tools are not only beneficial to examine neologisms. Other stylistic devices such as juxtapositions and metaphors could also be analyzed by using english-corpora.org. Hence, students could search for phrases like *heaven and hell* (juxtaposition), *eyes like* and *cold as* (both metaphors) and compare different types of texts (e.g., newspaper articles and fictional texts) to analyze the functions and effects of stylistic devices. Apart from letting your students work with english-corpora.org, you could also make use of the website's corpus tools to create more of your own corpus-based material. English-corpora.org contains several different corpora with various functions, such as the possibility to search in specific text genres, e.g., only fictional texts, in the Corpus of Contemporary American English (COCA). Thus, applying corpus methods to teach, learn and engage further with stylistic devices has great potential. Our ideas and proposed lesson plan are specifically designed to allow adjustments in order to fit your teaching unit and your individual learner group.

7 Caveats & limitations

Implementing ideas and creating new materials for EFL teaching may come with certain difficulties or obstacles. Our intention is to consider possible difficulties the designed lesson may cause as well as to mention the limitations of this chapter.

- Avoid swear words and inappropriate expressions that might come up in the search results.
- Make sure that all neologisms on the worksheet can actually be found in the corpus and that there are enough occurrences to find out the meaning of the specific neologism. Keep different proficiency levels of your students in mind (some students might need more support when working with the corpus-tool, especially if it is their first time working with it).
- It may be necessary to provide explanations of possibly unfamiliar functional terms such as “collocates”

or “lemmas”.

Lastly, the corpus tool itself holds one small limitation regarding the number of searches you want to perform. After a few searches, english-corpora.org will ask you to upgrade to a premium account (Fig. 13). You do not need to update: you can simply wait and then click on “Click here to continue with your search” when it appears (Fig. 13).

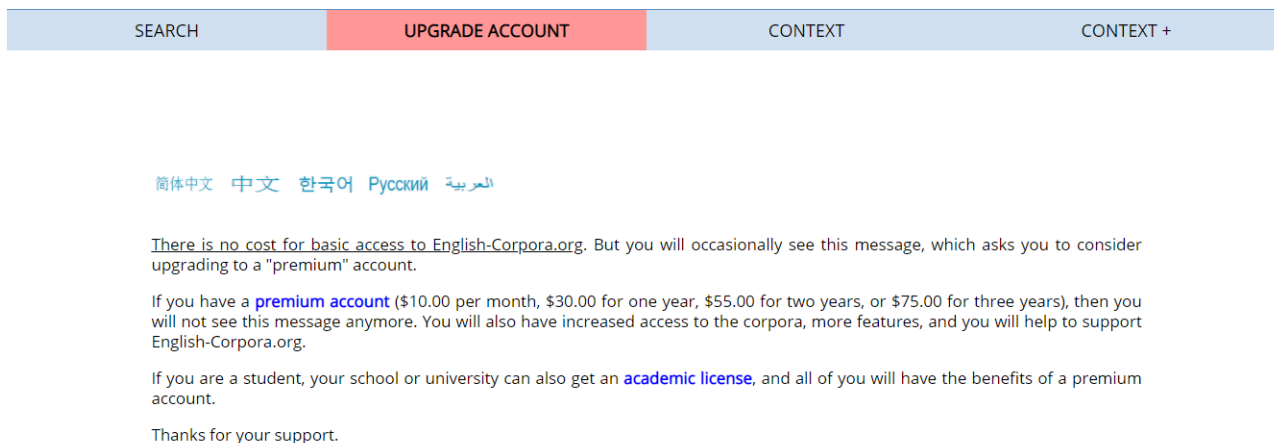


Fig. 13: Premium account notification

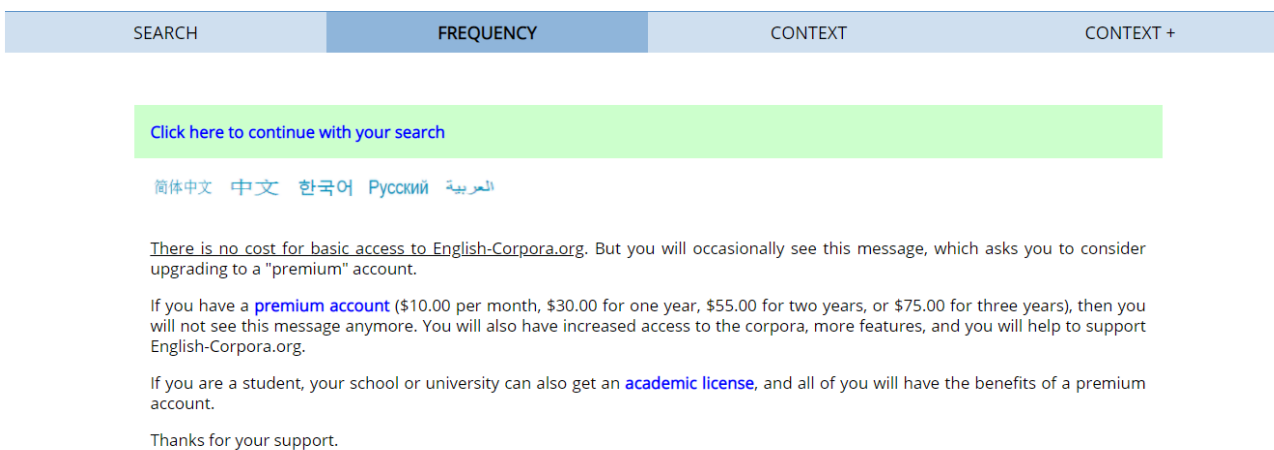


Fig. 14: Continuing your search

8 Conclusion

The method of using corpus tools in the EFL classroom might be quite new to teachers but it can provide promising opportunities that promote progressive and innovative teaching while at the same time being compatible with the curriculum. After this lesson, students and teachers should be familiar with basic features and functions of the NOW corpus and english-corpora.org in general, which is a great foundation for employing corpus-informed language teaching materials in further lessons. Corpus work can constitute a valuable benefit because it improves explicit learning and encourages students to work autonomously. The choice of topic and methods has great potential to activate and motivate students, as unfamiliar, possibly funny-sounding words may spark students' interest in learning more about them, thus keeping them focused on the lesson. Furthermore, students become familiar with more techniques to facilitate intelligent guessing and figuring out what a word means and in which contexts it can be used without

using a dictionary. Finally, all the ideas and thought-provoking impulses may encourage more teachers to employ new and more diversified teaching materials in their classes, as it can lead to enriching and motivating lessons in the EFL classroom.

9 Resources and references

Cambridge Advanced Learner's Dictionary. 2021. Butterfingers". <https://dictionary.cambridge.org/us/dictionary/english/butterfingers> (29 March, 2021).

Corpus of Contemporary American English (COCA). <https://www.english-corpora.org/coca/>.

Electoral Commission. 2016. Referendum on the UK's Membership of the European Union. <http://web.archive.org/web/20160630154529/http://www.electoralcommission.org.uk/find-information-by-subject/elections-and-referendums/upcoming-elections-and-referendums/eu-referendum> (29 March, 2021).

English-corpora.org. www.english-corpora.org/.

Ewing Bölke, Susannah. 2020. Brand New Neologisms. *English Explorations*. <https://englishexplorations.check.uni-hamburg.de/brand-new-neologisms/> (16 May, 2021).

Locke, Gary. 2020. Neologisms, 2020 Edition. *English Language Blog*. <https://blogs.transparent.com/english/neologisms-2020-edition/> (16 May, 2021).

News on the Web Corpus (NOW). <https://www.english-corpora.org/now/>.

Niedersächsisches Kultusministerium. 2015. Kerncurriculum für das Gymnasium, Schuljahrgänge 5-10: Englisch. https://cuvo.nibis.de/cuvo.php?k0_0=Dokumentenart&v0_0=Kerncurriculum&k0_1=Schulbereich&v0_1=Sek+I&k0_2=Schulform&v0_2=alle&k0_3=Fach&v0_3=Englisch&docid=1197&p=detail_view (28 June, 2021).

Niedersächsisches Kultusministerium. 2017. Kerncurriculum für das Gymnasium, gymnasiale Oberstufe: Englisch. https://cuvo.nibis.de/cuvo.php?k0_0=Schulbereich&v0_0=Sek+II&k0_1=Fach&v0_1=Englisch&docid=1132&p=detail_view (28 June, 2021).

Word Spy. www.wordspy.com/.

10 Appendix

Answer sheet: Corona neologisms and their meanings

Neologism	Meaning
Zoomer	Someone from Generation Z (the expression alludes to the term 'boomer', i.e., a member of the baby boomer generation).
Coronahobbies	A hobby discovered or more regularly practised due to Covid.
Doomscrolling	Reading through all the bad Corona related news on a technical device.
Zombie	Feeling like a zombie becoming after giving/attending classes on Zoom all week.
Coronacleaning	Cleaning as a hobby during the corona pandemic.
Procrastibaking	Baking as a hobby during the corona pandemic.
Hairpocalypse	A terrible haircut after a long time of not having been to the hairdresser.
Quarancut	A haircut given at home (during quarantine).
Covidiots	People who ignore the shutdown and hygiene measures of Covid.
Mask-shaming	Denouncing someone for not wearing or wearing a mask.
Quaranteen	A teenager who spends parts of their youth in shutdown/quarantine.
Blursday	A day on which one does not notice time passing by.

Student step-by-step guide

Student step-by-step guide

How to use English-corpora.org

Log in on english-corpora.org.

Select "Overview".

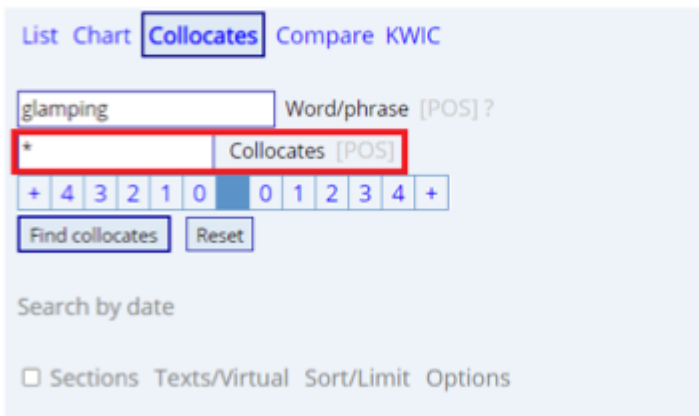
You now see a list of corpora. Choose the News on the Web (NOW) corpus.

Change your search settings from "List" to "Collocates".



The screenshot shows the search interface with the 'Collocates' tab highlighted. The search box is empty, and the 'Collocates [POS]' dropdown is also empty. The navigation bar includes 'List', 'Chart', 'Collocates', and 'Compare KWIC'. Below the search box is a row of buttons: '+ 4 3 2 1 0 0 1 2 3 4 +'. The 'Find collocates' and 'Reset' buttons are visible. The 'Search by date' section is empty, and the 'Sections Texts/Virtual Sort/Limit Options' section is also empty.

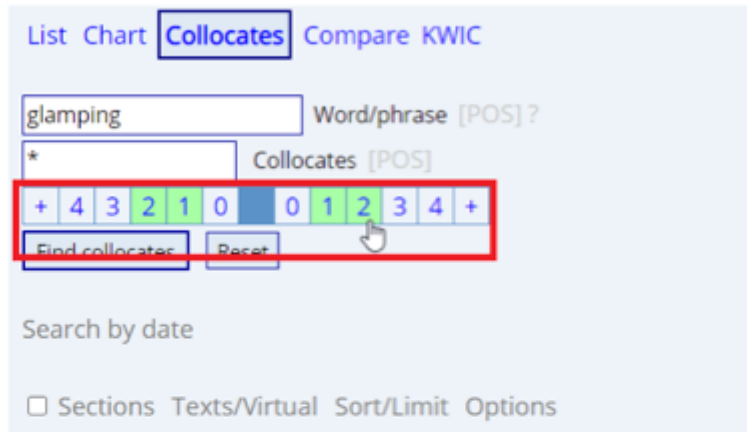
Type the word you want to investigate further into the search box labelled "Word/phrase". Select a particular POS (Part of Speech) or insert a star (*) to search for all word classes.



The screenshot shows the search interface with 'glamping' entered in the search box and '*' selected in the 'Collocates [POS]' dropdown. The navigation bar includes 'List', 'Chart', 'Collocates', and 'Compare KWIC'. Below the search box is a row of buttons: '+ 4 3 2 1 0 0 1 2 3 4 +'. The 'Find collocates' and 'Reset' buttons are visible. The 'Search by date' section is empty, and the 'Sections Texts/Virtual Sort/Limit Options' section is also empty.

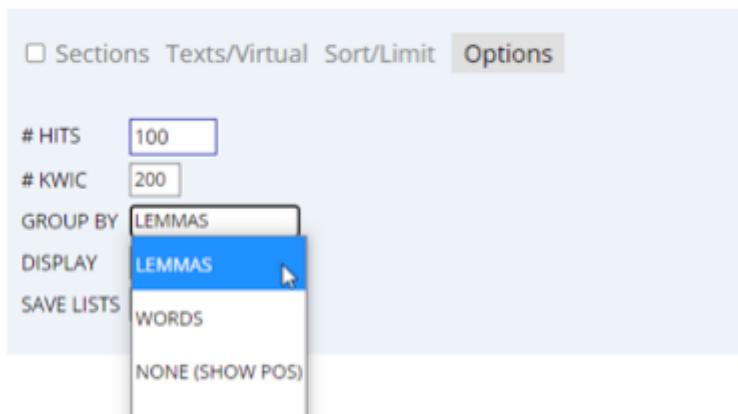
This worksheet was downloaded from: Winkelmann, Sarah & Nicol Sperling (2021). "Teaching neologisms using corpus tools." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

You can specify the distance between the analyzed word and its collocates with the help of the bar of numbers underneath. The colored box (dark blue) between the zeros represents the searchword; the numbers on the left represent the collocates preceding the word; the numbers on the right represent the collocates that follow the word.



Next up is adjusting “Options”.

Change “Group by Words” into “Group by Lemmas”.



Now click on “Find collocates” to start your search.

This worksheet was downloaded from: Winkelmann, Sarah & Nicol Sperling (2021). “Teaching neologisms using corpus tools.” In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

List Chart **Collocates** Compare KWIC

glamping Word/phrase [POS] ?

* Collocates [POS]

+ 4 3 2 1 0 0 1 2 3 4 +

Find collocates Reset

Search by date

Sections Texts/Virtual Sort/Limit **Options**

HITS

KWIC

GROUP BY

DISPLAY

SAVE LISTS

Select a collocate from the list to examine the meaning of the neologism.

ON CLICK: **CONTEXT** TRANSLATE (??) GOOGLE IMAGE PRON-VIDEO BOOK (HELP)

HELP			FREQ	PER MIL	
1	<input type="checkbox"/>	[THE]	644	0.00	
2	<input type="checkbox"/>	[A]	636	0.00	
3	<input type="checkbox"/>	[AND]	566	0.00	
4	<input type="checkbox"/>	[BE]	549	0.00	
5	<input type="checkbox"/>	[I]	518	0.00	
6	<input type="checkbox"/>	[IN]	490	0.00	
7	<input type="checkbox"/>	[SITE]	329	0.00	
8	<input type="checkbox"/>	[OF]	288	0.00	

When you click on the number next to the example, the interface will show you a larger excerpt of the article featuring the neologism and the chosen collocate.

CLICK FOR MORE CONTEXT EXPLORE NEW FEATURES SAVE TRANSLATE ANALYZE

1	21-03-27 CA	mountainviewtoday			barn renovated to accommodate gatherings and weddings. # The intention is to have each glamping site serviced with water, electricity and compo
2	21-03-27 GB	birminghammail.co.uk			holiday rentals. But it actually encompasses a far broader range, including lodges, glamping sites , caravans and other holiday accommodation that f
3	21-02-26 US	New York Times			year over year. Campspot, a platform for booking campgrounds and R.V. sites , glamping tents and cabins, said reservations are up 50 percent over l
4	21-02-26 HK	thestandard.com.hk			This would involve providing recreational elements such as additional lookout points, treetop adventure and glamping sites , better toilet facilities ar
5	21-02-25 HK	hk.asiatatler.com			, stating that recreational elements such as " additional lookout points, treetop adventure and glamping sites " will be added over the next two to th
6	21-02-25 HK	hk.asiatatler.com			, with numerous private glamping facilities booked out during holidays. # Aside from building glamping sites , the government is also planning ahead
7	21-02-24 HK	hongkongfp.com			government will be " providing recreational elements like additional lookout points, treetop adventure and glamping sites , improving toilet facilities
8	21-02-11 US	Architectural Digest			to be further along on their expansion plans, which include the construction of a glamping site . (Despite these efforts, much of the island, which as c

This worksheet was downloaded from: Winkelmann, Sarah & Nicol Sperling (2021). "Teaching neologisms using corpus tools." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

Source information:

Source	https://www.mountainviewtoday.ca/mountain-view-county-news/glamping-retreat-proposed-for-water-valley-3546856
Country/date	CA (21-03-27)
Title	Glamping retreat proposed for Water Valley

Expanded context:

that enters the property from the west, connects to a man-made lake, and exists the property on its eastern boundary at Rgs. Rd. 52. # The eastern and southern portions of the parcel are forested, and the western and northwestern portions are on the site of an historic tree farm. # The site is currently zoned residential low density. The plan is to have the property rezoned direct control district. # The plan calls for 50 accommodation tents, with the existing barn renovated to accommodate gatherings and weddings. # The intention is to have each **glamping site** serviced with water, electricity and composting toilets. # "That is the luxury experience we are going for," he said. "We want to be environmentally friendly." # The project will have three development phases, with the first involving the development of the northeastern access point and accommodation along the north side of the lake, and will included renovation of the barn into a wedding and reception centre. # FireSmart and BearSmart plans would be put in place, he said, noting the

Use the navigation bar at the top of the website.



Go back to "Frequency" and look at other collocates to confirm your theories on the meaning of the neologism.

This worksheet was downloaded from: Winkelmann, Sarah & Nicol Sperling (2021). "Teaching neologisms using corpus tools." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

8. Commas in argumentative writing

MARIE NOTTMEIER, ALINA SOPHIE PETERS, AND LARA WARNECKE

1 Introduction and rationale

Following a corpus-based lesson structure, the proposed lesson enables students to deduce frequent comma placement in a specified context (here: academic writing). Thereby, students and teachers both interact with the English language via the dedicated corpus and corpus tool (here: BNC1994 on Sketch Engine), drawing their own insights and gaining sustainable strategic competences (Niedersächsisches Kultusministerium 2015: 8), which we believe enhances motivation and successful learning in the long run. The overarching aim is to develop independent, autonomous learning strategies and agency. This is supported with various scaffolding options.

Outline and Objectives

Target learners: Secondary school (10th grade or higher); B1/B2 level students

Teacher preparation time (multiple options):

1. Use the worksheets included in this chapter (see Appendix)
 - Getting familiar with the proposed lesson structure: 15 min
2. Collect your own data and design worksheet(s)
 - Getting familiar with Sketch Engine: 10 min
 - Creating concordance lines and worksheet(s): 30 min
 - Optional: worksheet with scaffolding: approx. 15 min

Lesson time: 90 min

Global learning objective: Students can use corpus extracts to deduce rules for the placement of commas and can apply them in their own argumentative essays.

Competence-oriented learning objectives:

Students can...

- spot grammatical patterns in concordance lines taken from a corpus (e.g., commas after introductory phrases, after conjunctive adverbs, and in enumerations; cf. 3.2.1).
- infer grammatical rules for the placement of commas based on the deduced patterns.
- apply some of these patterns in their own argumentative essays.

2 Corpus, tool and methods

This chapter draws on the British National Corpus 1994 (hereafter BNC1994), which consists of approximately 100 million words of native British English. The BNC1994 is a balanced corpus, which, among other registers, features written academic language.

The planned lesson makes use of the online corpus tool Sketch Engine. Sketch Engine grants a free 30-day trial and is free of charge until March 2022 for researchers, lecturers, and students from academic institutions in the EU participating in the ELEXIS project.

The lesson at hand takes a direct corpus approach: through teacher-corpus interaction, appropriate concordance lines are selected and presented in two worksheets. The focus on simple and easy-to-use functions of the corpus tool facilitates preparation for teachers, as they do not necessarily need to have any experience in using corpora and do not have to spend a lot of time familiarising themselves with the various corpus tool functions.

In this form of data-driven learning, learners do not search in corpora themselves, but engage in paper-based corpus interaction. In contrast to an indirect use, in which the corpus only has an impact on the material, students work and interact directly with the concordance lines. This can help students to get accustomed to concordance lines and should facilitate future lessons in which students consult the corpora themselves.

3 Step-by-step guide

3.1 Teacher preparation

3.1.1 Corpus queries – data collection

This approach features teacher-corpus as well as student-corpus interaction. As the students' interaction will be paper-based and consist of simplified corpus results, the lesson requires some teacher preparation. As part of this preparation, teachers will engage with the corpus, conduct corpus queries and collect data that will form the basis for the students' paper-based corpus interaction and their worksheet. To compile the data, follow the steps described in this chapter.

Steps:

1. Visit www.sketchengine.eu. This is our corpus tool.
2. In order to use Sketch Engine, you will need to create an account. By clicking on “FREE Trial” at the top of the page (Fig. 1, a), you can sign up for a free 30-day trial, which enables you to use all the search-related functions needed for our purpose.



Fig. 1: Home page Sketch Engine

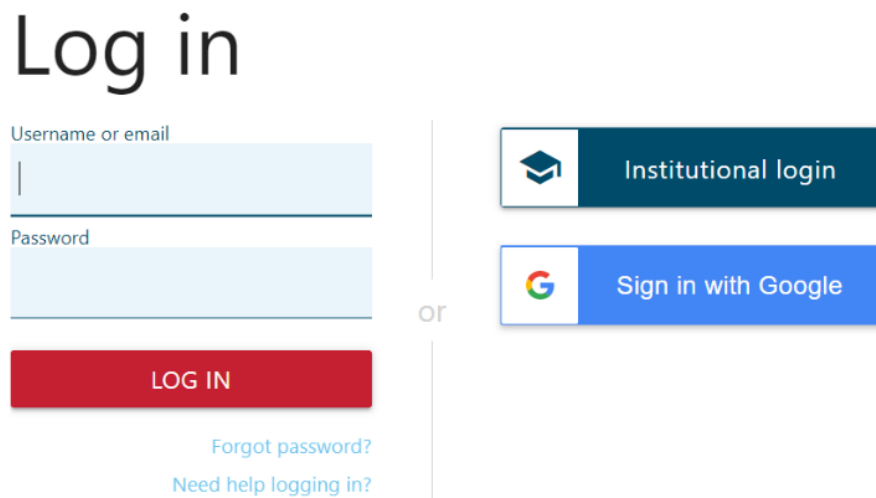


Fig. 2: Log-in interface

3. Then, click on “LOG IN” (Fig. 1, b). An interface (Fig. 2) will pop up, in which you can type the username (or email) and password you created in step 2). Press “LOG IN” to log into your account.
4. You should now be able to see the following dashboard (Fig. 3). Next to the interface for selecting a corpus, there is an embedded video on the right-hand side. By clicking on it, you will see a short and very informative introduction to Sketch Engine. It is not necessary to watch it when following this step-by-step guide. However, as this quick start tutorial gives you background information on the corpus tool, it might be helpful for you when modifying or creating your own lesson.

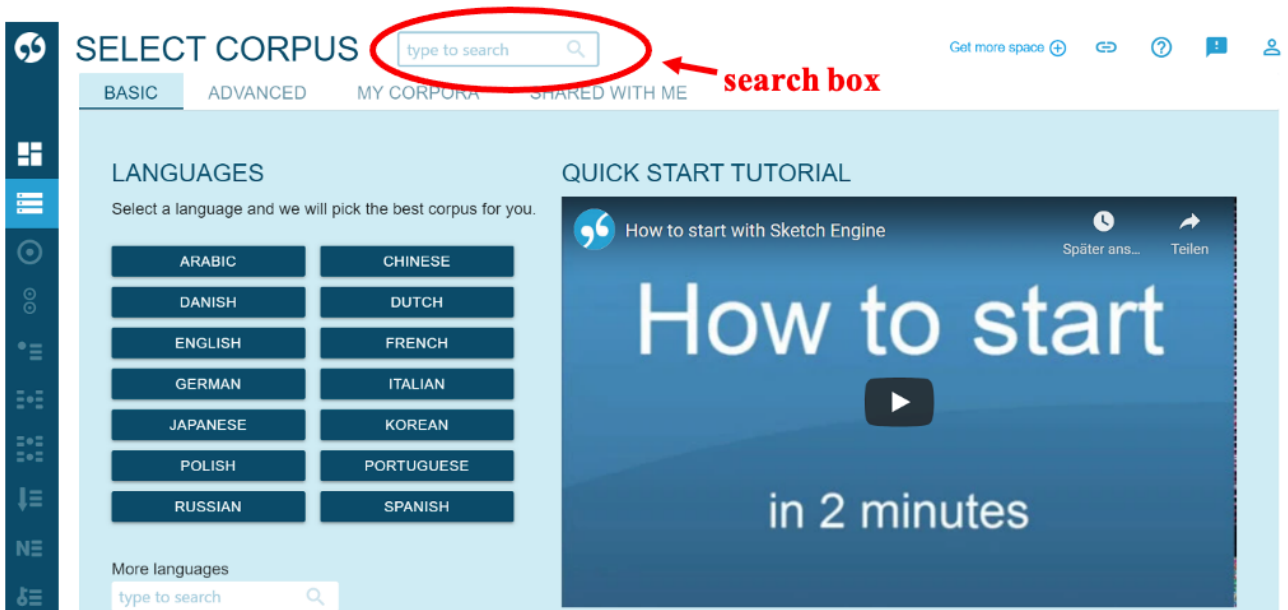


Fig. 3: Interface for selecting a corpus

5. By clicking onto the search box at the top (Fig. 3), a specific corpus can be selected. As explained above, we will use the BNC, so please type "BNC" into the search box.
6. In the appearing drop-down menu, select "British National Corpus (BNC)" by clicking on it (Fig. 4).

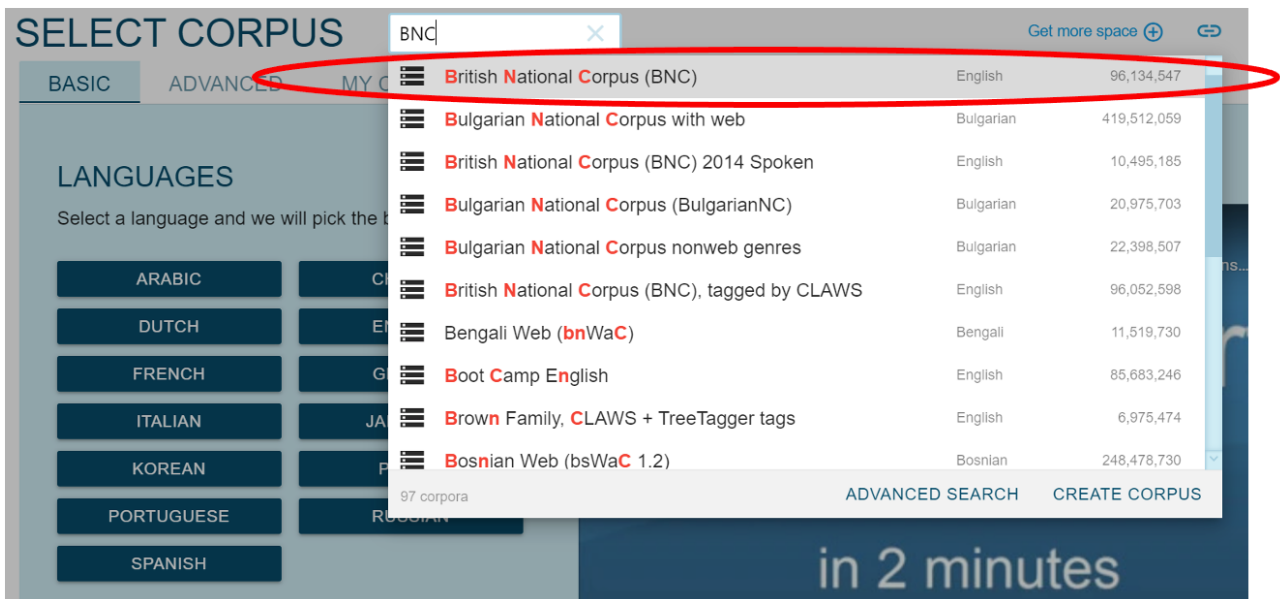


Fig. 4: Copus selection

7. The interface should now look like the one in Fig. 5. On this interface, you get an overview of all the different functions that can be used. If you are interested in getting detailed information about the chosen corpus, such as the number of words or the kind of text genres it encompasses, you can click on "CORPUS INFO" (Fig. 5, a).
8. As we want to deduce patterns for the placement of commas from entire sentences, we need to use the "Concordance" function (Fig. 5, b), so please click on it.

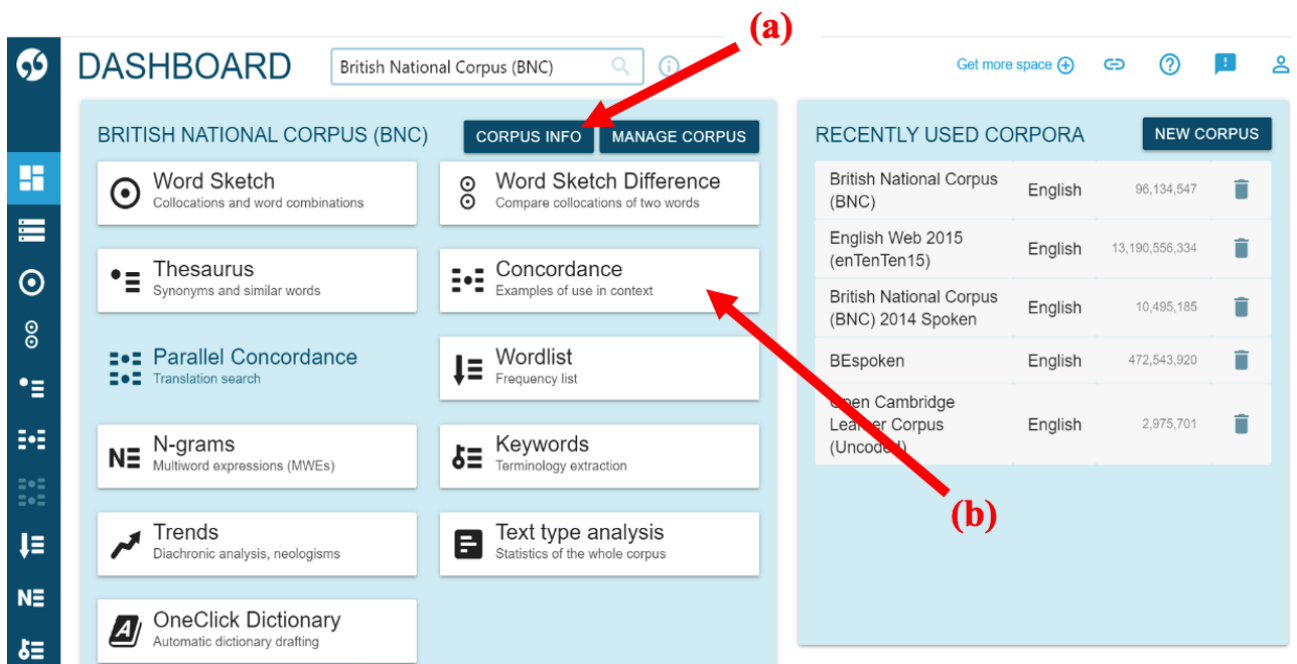


Fig. 5: Function overview

- The interface will change again into the dashboard shown in Fig. 6. Click on the tab labelled "ADVANCED", which enables us to specify a register.



Fig. 6: Search interface

- The drop-down menu called "Query Type" should already be open on the next interface. Now choose the option "simple" (Fig. 7, a). This option is not case sensitive and finds the search word form of your choice.

CONCORDANCE

British National Corpus (BNC)



BASIC

ADVANCED

ABOUT

Query type ?

- simple (a)
- lemma
- phrase
- word
- character
- CQL

Subcorpus ?

none (the whole corpus) (b)

Filter context ? v

Text types ? v

Fig. 7: Advanced search interface

11. Then, choose a subcorpus. To do so, click on the arrow next to “none (the whole corpus)” (Fig. 7, b). In the appearing drop-down menu, pick “Written Academic” (Fig. 8). This will limit our search to the selected register.

- none (the whole corpus)
- 1960-1974
- 1975-1984
- 1985-1993
- Audio sentences mp3
- Spoken transcripts
- Written Academic
- Written Domain Imaginative
- Written Domain Informative
- Written Medium Book

Fig. 8: Subcorpus selection

12. Type a comma (“,”) into the search field which says “abc” (Fig. 9, a).

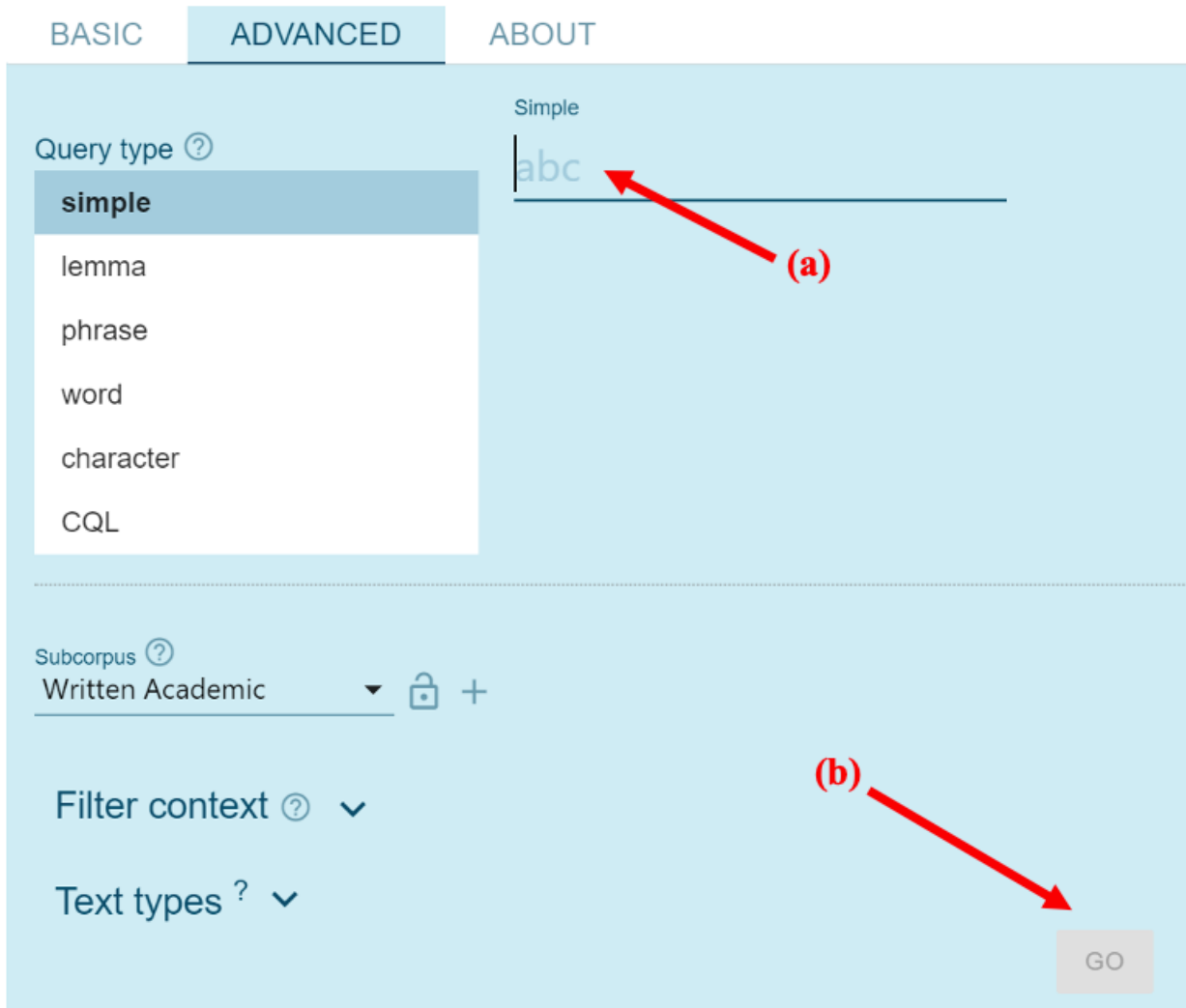


Fig. 9: Starting the search

13. Now, click on “GO” (Fig. 9, b) in order to start the search. (Note: the box labelled “GO” turns red once you type something into the search field.)
14. You should now see your search results (Fig. 10). To get a random sample, click on the dice (Fig. 10, a). This mixes up the order of the concordance lines and prevents us from seeing only sentences from the same source. By doing that, the authenticity of the output is increased. In the pop-up box (Fig. 11), we can select the sample size. A sample of 200 concordance lines is preselected, but for our purposes, we will change it to 100 in order to reduce the number of concordance lines to a more manageable number.

CONCORDANCE British National Corpus (BNC)

Written Academic simple • 747,515 42,407.19 per million tokens • 0.67%

Details Left context KWIC Right context

1	<input type="checkbox"/>	<i>Written books a... (1979) and Brilliant (1974).</i>	,	recommendations of appropriate taste or choice of style are
2	<input type="checkbox"/>	<i>Written books a... mainly concerned with the architectural aspects of mosaic</i>	,	e.g. the technicalities of foundations.
3	<input type="checkbox"/>	<i>Written books a... of foundations.</i>	,	meanwhile, echoes much of that discussed by Vitruvius - w
4	<input type="checkbox"/>	<i>Written books a... r.</i>	,	echoes much of that discussed by Vitruvius - with only slight
5	<input type="checkbox"/>	<i>Written books a... of that discussed by Vitruvius - with only slight differences</i>	,	e.g. the types of pavements to which particular foundations
6	<input type="checkbox"/>	<i>Written books a... new light on the broader "aesthetic environment" of mosaic</i>	,	such as Lucian's Dream, are of little assistance.
7	<input type="checkbox"/>	<i>Written books a... aesthetic environment" of mosaic, such as Lucian's Dream</i>	,	are of little assistance.
8	<input type="checkbox"/>	<i>Written books a... ucian's Dream, are of little assistance.</i>	,	within mosaics and elsewhere, provide a richer source of in
9	<input type="checkbox"/>	<i>Written books a... stance.</i>	,	provide a richer source of information.
10	<input type="checkbox"/>	<i>Written books a... n.</i>	,	but it is clear that under some conditions such remains can
11	<input type="checkbox"/>	<i>Written books a... nditions such remains can suggest the name of a sponsor</i>	,	the reason for sponsorship, and the presence of a respecte
12	<input type="checkbox"/>	<i>Written books a... suggest the name of a sponsor, the reason for sponsorship</i>	,	and the presence of a respected craftsmen or group of craf
13	<input type="checkbox"/>	<i>Written books a... craftsmen At Avenches - and probably at Bavay - in France</i>	,	and at Seriana and Timgad in North Africa, the names of in

Fig. 10: Randomising the sample

GET A RANDOM SAMPLE

When working with a large concordance, random sample is used to reduce the number of concordance lines while preserving the representativeness of the sample. The number of lines defined by the user will be randomly selected from all parts of the corpus. The lines will stay in the same order as they appear in the corpus.

A random sample with the same number of lines generated from the same concordance will always produce **exactly the same** concordance lines. This behaviour is intentional so that different users (e.g. students) can follow the same steps and arrive at the same result. To generate a different sample, a different number should be used. Setting the value to 201 instead of 200 will generate a completely different sample.

200 GO

Fig. 11: Interface to create a random sample

- Now, you can see a randomly sampled list of concordance lines that have a comma in it (Fig. 12). By clicking on the arrow next to "KWIC" (Fig. 12), you can change the display mode. There are two display options: the first one, "KWIC", shows the comma in the centre and some context to the left and right; the second one, "sentence", shows the complete sentence. We chose to select "sentence", as this makes it easier to filter out concordance lines that are not appropriate to the language level of the students.

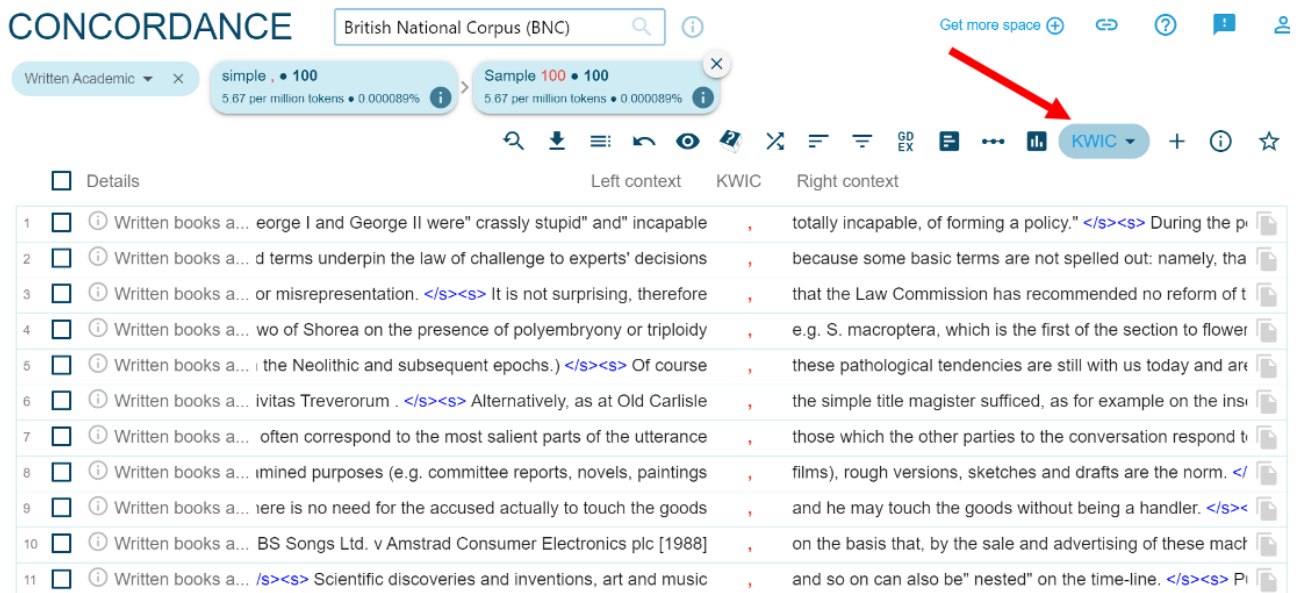


Fig. 12: Change display mode

16. By ticking the boxes next to the individual concordance lines, you can select the lines you want to use for your worksheet and filter out those that do not fit the purpose or the language level (Fig. 13, a). For example, in line 4 there are many specific technical terms which the students might not understand. As this lesson's focus is on punctuation rather than content, it would be counterproductive to include this line, so we did not tick the box next to it. Confirm your selection by clicking on "SELECTED" beneath "show only" (Fig. 13, b).

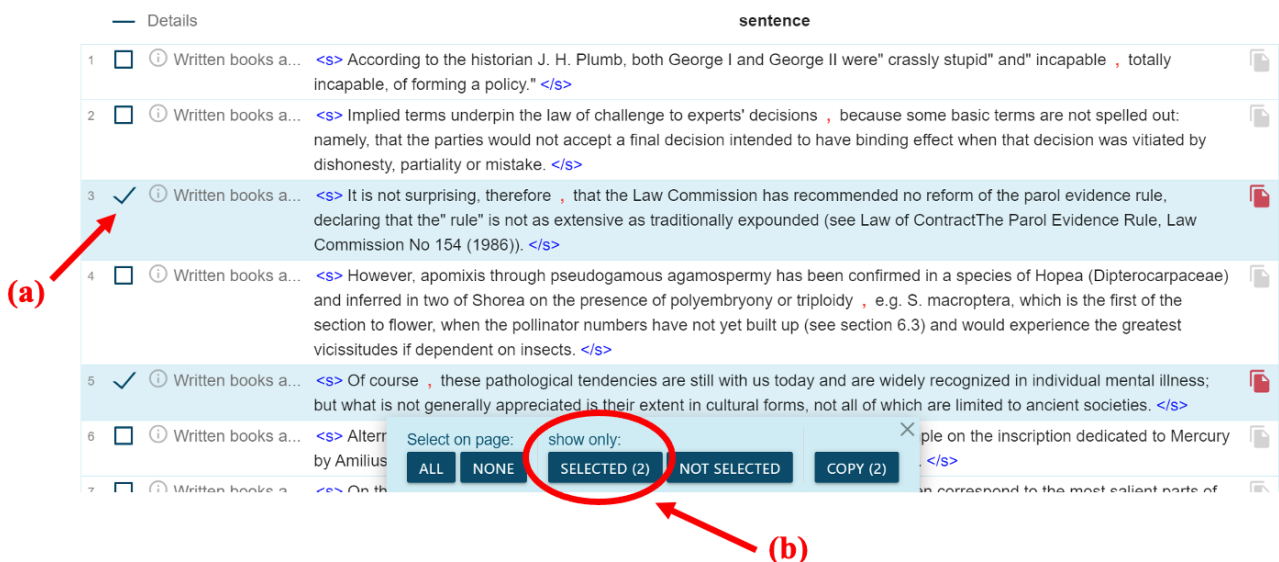


Fig. 13: Selecting concordance lines

17. Now, you have a selected list of concordance lines appropriate for the context. This list will provide the basis for the worksheet. You can download it by clicking on the download symbol (Fig. 14).

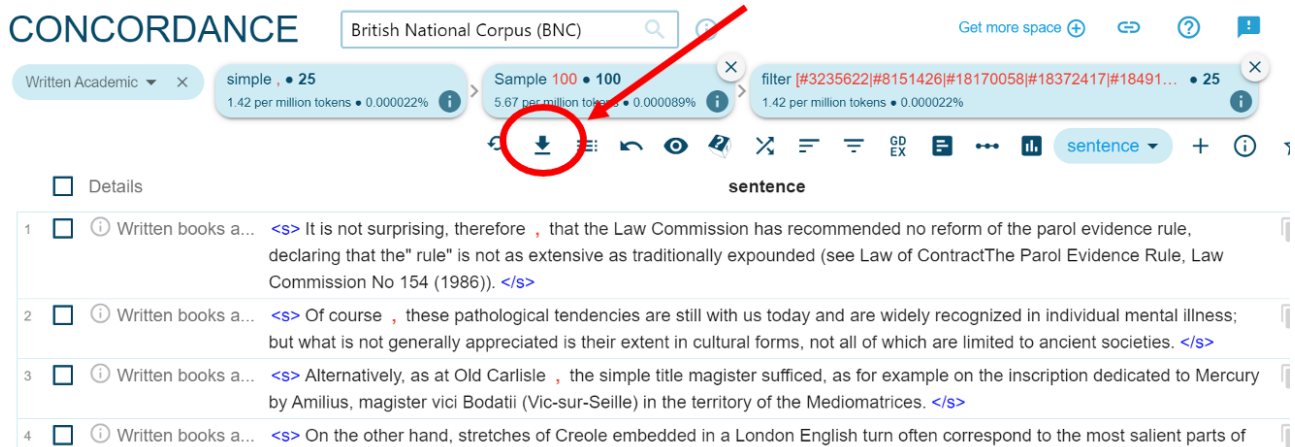


Fig. 14: Downloading the results

18. A pop-up window will open. Click on “PDF” to save the list of concordance lines as a PDF document (Fig. 15). We recommend this format, as the layout is very clear and easy to work with.

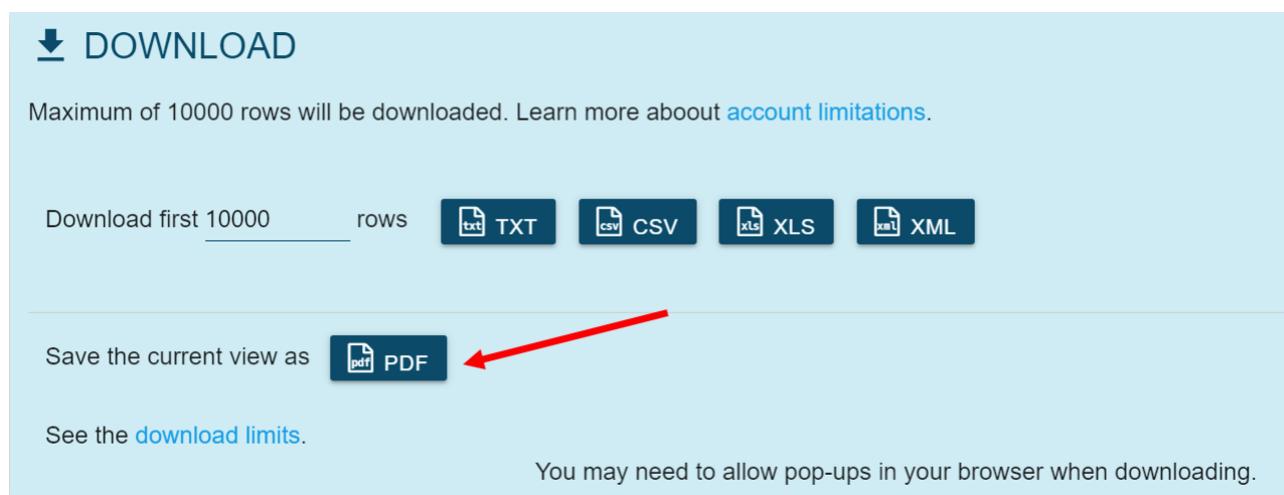


Fig. 15: Download interface

19. Now you have successfully collected data that your students can use in order to deduce patterns for the placement of commas.

How to create frequency lists

Frequency lists provide an overview of the words that occur most frequently directly before and after a comma. Such lists can help to empirically validate the patterns deduced by the students. Once you have reached step 13 (see above), you may execute the following steps in order to generate your own frequency lists.

Considering step 13 above as a starting point, you may now click on the “Frequency” symbol, which is displayed in the header right above the concordance lines (Fig. 16).

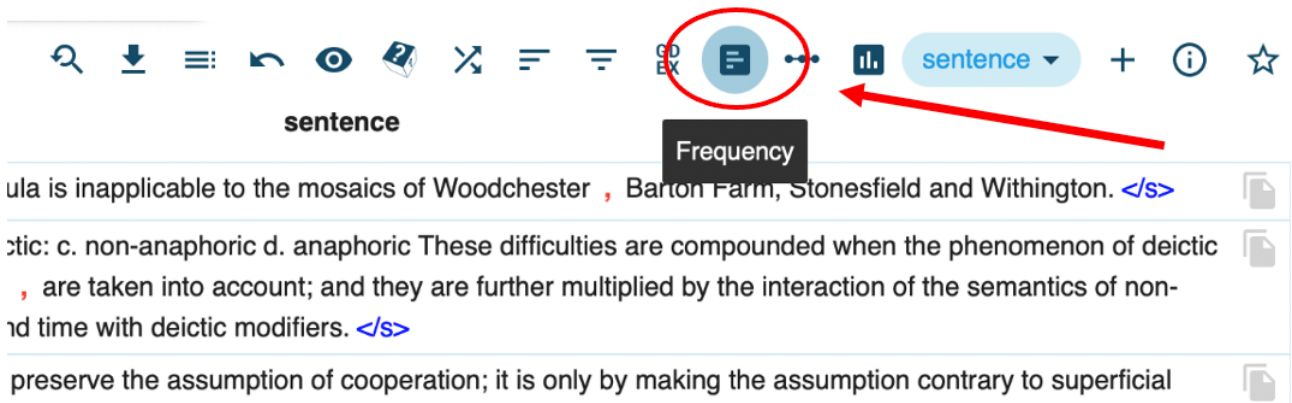


Fig. 16: Generating frequency lists

In the interface below, choose the option “word” as an attribute and click on “1” on the left-hand side of “KWIC” (Fig. 17). This will generate a list of words which are often followed by a comma, sorted by frequency. Click “GO” to start your search.



Fig. 17: Interface for generating frequency lists

You should now be able to see a frequency list similar to the one in Fig. 18. Depending on the corpus you consulted, the generated frequency list may vary. You can either screenshot and save your results or tick the boxes and download them (cf. steps 16-18).

	<input type="checkbox"/>	Word	Frequency ↓	Relative ?		
1	<input type="checkbox"/>)	18,084	160.97		...
2	<input type="checkbox"/>	'	9,244	82.28		...
3	<input type="checkbox"/>	however	8,861	78.87		...
4	<input type="checkbox"/>	example	7,795	69.38		...
5	<input type="checkbox"/>	However	7,743	68.92		...
6	<input type="checkbox"/>	is	4,731	42.11		...
7	<input type="checkbox"/>	and	4,620	41.12		...
8	<input type="checkbox"/>	that	4,246	37.79		...
9	<input type="checkbox"/>	it	2,850	25.37		...
10	<input type="checkbox"/>	course	2,558	22.77		...
11	<input type="checkbox"/>	time	2,346	20.88		...
12	<input type="checkbox"/>	"	2,222	19.78		...
13	<input type="checkbox"/>	instance	2,135	19.00		...
14	<input type="checkbox"/>	therefore	2,084	18.55		...
15	<input type="checkbox"/>	case	2,022	18.00		...
16	<input type="checkbox"/>	hand	1,993	17.74		...
17	<input type="checkbox"/>	which	1,921	17.10		...
18	<input type="checkbox"/>	not	1,913	17.03		...
19	<input type="checkbox"/>	them	1,822	16.22		...
20	<input type="checkbox"/>	Thus	1,711	15.23		...
21	<input type="checkbox"/>	this	1,634	14.54		...
22	<input type="checkbox"/>	then	1,600	14.24		...
23	<input type="checkbox"/>	Indeed	1,476	13.14		...
24	<input type="checkbox"/>	Moreover	1,474	13.12		...

Fig. 18: Frequency list – words followed by a comma

If you click “CHANGE CRITERIA”, you will be taken back to the frequency interface.

Frequency

CHANGE CRITERIA

BACK TO CONCORDANCE

Fig. 18: Return to the interface

Now, to obtain a list of words that frequently appear after a comma, select “1” on the right hand-side of “KWIC”, as displayed in Fig. 19. Then, execute your search via the “GO” button.

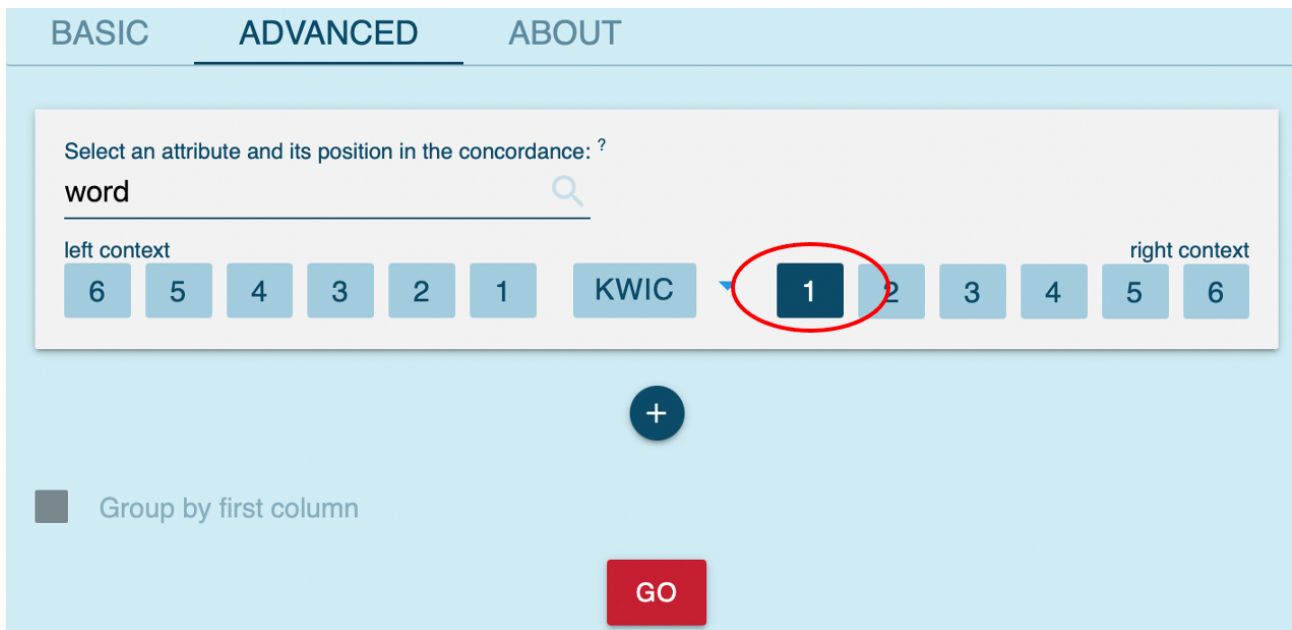


Fig. 19: Interface for generating frequency lists

Once again, either screenshot and save your results or tick the boxes and download them (cf. steps 16-18).

		Word	Frequency ↓	Relative ?		
1	<input type="checkbox"/>	and	75,005	667.63		...
2	<input type="checkbox"/>	the	55,369	492.84		...
3	<input type="checkbox"/>	but	25,673	228.52		...
4	<input type="checkbox"/>	in	19,505	173.62		...
5	<input type="checkbox"/>	it	15,912	141.63		...
6	<input type="checkbox"/>	as	15,911	141.63		...
7	<input type="checkbox"/>	which	14,948	133.05		...
8	<input type="checkbox"/>	or	14,053	125.09		...
9	<input type="checkbox"/>	a	12,955	115.31		...
10	<input type="checkbox"/>	for	11,977	106.61		...
11	<input type="checkbox"/>	however	9,203	81.92		...
12	<input type="checkbox"/>	to	8,553	76.13		...
13	<input type="checkbox"/>	that	8,351	74.33		...
14	<input type="checkbox"/>	is	7,600	67.65		...
15	<input type="checkbox"/>	with	7,333	65.27		...
16	<input type="checkbox"/>	there	6,719	59.81		...
17	<input type="checkbox"/>	'	6,056	53.91		...
18	<input type="checkbox"/>	if	5,440	48.42		...
19	<input type="checkbox"/>	we	5,412	48.17		...
20	<input type="checkbox"/>	of	5,262	46.84		...
21	<input type="checkbox"/>	they	5,075	45.17		...
22	<input type="checkbox"/>	he	4,551	40.51		...
23	<input type="checkbox"/>	then	4,534	40.36		...
24	<input type="checkbox"/>	although	4,506	40.11		...

Fig. 20: Frequency list – words preceded by a comma

3.1.2 Worksheet and task creation

For the first task, you need to design a worksheet based on the corpus data you compiled in section 3.1.1. As the students are supposed to examine the concordance lines and deduce patterns themselves, the worksheet basically consists of the raw corpus results (though filtered by the teacher in section 3.1.1). You can highlight the comma in each sentence in order to draw the students' attention to that specific part of the sentence. Add a task description. Your worksheet could look like the sample worksheet in Fig. 21. You could also use the sample worksheet in Fig. 21: simply click on it to download the PDF version.

Commas in Argumentative Writing I

In the following table, you can see sentences containing one or more commas. The sentences, or so-called concordance lines, have been collected using the British National Corpus. This means that these sentences stem from a range of different academic texts that present authentic language use.

Tasks:

1. a) Examine the concordance lines and identify patterns of comma placement in academic writing. In each sentence, the comma in question is highlighted.
- b) Deduce comma rules from the sentences.

(Hint: Are there specific words or constructions that require a comma? Where exactly is the comma placed?)

Details	sentence
1	Written books a. ... <S> It is not surprising, therefore, that the Law Commission has recommended no reform of the parole evidence rule, declaring that the "rule" is not as extensive as traditionally expounded (see Law of Contract/The Parole Evidence Rule, Law Commission No 154 (1986)). </S>
2	Written books a. ... <S> Of course, these pathological tendencies are still with us today and are widely recognized in individual mental illness; but what is not generally appreciated is their extent in cultural forms, not all of which are limited to ancient societies. </S>
3	Written books a. ... <S> Alternatively, as at Old Carlisle, the simple title magister sufficed, as for example on the inscription dedicated to Mercury by Amilius, magister vici Bodati (Vic-sur-Seille) in the territory of the Mediomatrici. </S>
4	Written books a. ... <S> On the other hand, stretches of Creole embedded in a London English turn often correspond to the most salient parts of the utterance, those which the other parties to the conversation respond to. </S>
5	Written books a. ... <S> Remember that in the production of texts for non-examined purposes (e.g. committee reports, novels, paintings, films), rough versions, sketches and drafts are the norm. </S>
6	Written books a. ... <S> Scientific discoveries and inventions in art and music, and so on can also be "nested" on the time-line. </S>
7	Written books a. ... <S> Having run through the preventive health catechism (more exercise, less tobacco and alcohol, and better diet) the 1979 report bluntly dismissed the idea of replacing Britain's tax-funded health service with some form of insurance scheme. </S>
8	Written books a. ... <S> In the words of Sallis again: "If we know that the head's motive is to protect the school, and especially the more needy in it, from questions which are ill-informed, destructive or selfish, we have a clear objective which is to make sure we don't deserve or unwittingly attract such accusation (1988:17). </S>
9	Written books a. ... <S> This indicates an acceptance of the state structures, or a recognition of a need for the separation of Northern Ireland from the remaining state structure of Ireland, but at the same time a continuing opposition to Protestant - loyalist dominance. </S>
10	Written books a. ... <S> Countless examples can be studied in living creatures: seaboard harbours, great castles and manor houses, pots, bottles, soccer goalposts and pinball machines. </S>
11	Written books a. ... <S> The original position seems to presuppose not just a neutral theory of the good, but a liberal, individualistic conception according to which the best that can be wished for someone is the unimpeded pursuit of his own path, provided it does not interfere with the rights of others. </S>
12	Written books a. ... <S> Otherwise they were on their own, playing, but not playing a game for there was no structured game to play. </S>
13	Written books a. ... <S> This defence of Marxism carries over, however, to a critique of Modernism itself: not merely for the fact that Nietzsche attaches 'importance' to art but for the further claim ascribed to him, and therefore to Modernism itself, that the nature of aesthetic experience contains in itself the form of understanding proper to the world itself (p. 66). </S>

Fig. 21: Sample worksheet for the first task

Depending on your students' proficiency, you might want to create a second version of the worksheet that is easier and provides scaffolding. In order to adapt the worksheet to a lower proficiency level, group and arrange the concordance lines according to the pattern they show. That way, you can create blocks with two to three concordance lines and have the students deduce one pattern per block. This gives them more time to focus on the pattern itself rather than spending time searching for concordance lines that go together. In Fig. 22, you can see a sample version of such a worksheet that provides this kind of scaffolding. In addition to the prearranged blocks, this worksheet also has a reduced number of concordance lines to work through. You can find the full version of this worksheet in the appendix, too.

14	Written books a. ... <S> So things such as ourselves, which perceive and have conscious awareness, do so because they are (or have) minds. </S>
15	Written books a. ... <S> However, this leaves open the possibility that the Earth, Venus and Mars acquired most of their volatiles from volatile-rich bodies by post-accretionary winnowing. </S>
16	Written books a. ... <S> Moreover, this possibility is strengthened by the neglect of surface repositories, because regardless of how the various volatiles were acquired it is likely that a smaller fraction of the inert gases are retained in such reservoirs than are carbon and nitrogen. </S>
17	Written books a. ... <S> Certainly, it is important to study bureaucracies as institutions in their own right, even if we would not go as far as the poet Alexander Pope who wrote: </S>
18	Written books a. ... <S> Richards holds that in poetry the function of "feeling" tends to dominate that of "sense", while Jakobson identifies a special "poetic" function which can be found in many uses of language, but which dominates over other functions in poetry. </S>
19	Written books a. ... <S> This, the court said, was an irrelevant consideration which rendered his decision unlawful. </S>
20	Written books a. ... <S> The gap between them, according to Allhusse, is mediated by art. </S>
21	Written books a. ... <S> However there are a number of important controlling factors: firstly, the subject fixates centrally (usually) he has to report a digit which is presented in order to confirm this central fixation); secondly, the stimulus is exposed for less than 200 milliseconds to avoid eye movement; thirdly, the visual angle of the stimuli from the fixation point should be between 2.5' and 5' (otherwise, confounding effects occur). </S>
22	Written books a. ... <S> Furthermore, Rossi and Rosadini (1967), who also used the Amytal technique, reported that only one patient out of 74 right handers, all of whom received injections on both sides, did not have exclusively left hemisphere speech dominance. </S>
23	Written books a. ... <S> In that case, the opportunity to earn remuneration arose when the appellant obtained the grant of a tenancy of a public house owned by and tied to a large brewery company. </S>
24	Written books a. ... <S> Now if the prediction that follows from this more of premises turns out to be false (in our example, if the planet does not appear at the predicted location), then all that the logic of the situation permits us to conclude is that at least one of the premises must be false. </S>
25	Written books a. ... <S> If there are exclusively private events, in the sense that they are in principle, and not just empirically, inaccessible to more than one observer, it might be argued that such events could not be intelligibly claimed, let alone shown to be, subject to any laws, and this means that no rational explanatory model could be constructed for them. </S>

Use this part of the worksheet to formulate rules:

Commas in Argumentative Writing I (with scaffolding)

In the following table blocks, you can see sentences containing one or more commas. The sentences, or so-called concordance lines, have been collected using British National Corpus. This means that these sentences stem from a range of different academic texts that present authentic language use.

Tasks:

1. a) Examine the concordance lines and identify patterns of comma placement in academic writing. In each sentence, the comma in question is highlighted. You can find one pattern in the sentences in each block.

b) Deduce comma rules from the sentences in each block. (*Hint*: Are there specific words or constructions that require a comma? Where exactly is the comma placed?)

Block A

However , this leaves open the possibility that the Earth, Venus and Mars acquired most of their volatiles from volatile-rich bodies by post-accretional veneering.

Moreover , this possibility is strengthened by the neglect of surface repositories, because regardless of how the various volatiles were acquired it is likely that a smaller fraction of the inert gases are retained in such reservoirs than are carbon and nitrogen.

Furthermore , Rossi and Rosadini (1967), who also used the Amytal technique, reported that only one patient out of 74 right handers , all of whom received injections on both sides, did not have exclusively left hemisphere speech dominance.

Pattern/rule:

Block B

It is not surprising , therefore , that the Law Commission has recommended no reform of the parol evidence rule, declaring that the " rule" is not as extensive as traditionally expounded (see Law of Contract The Parol Evidence Rule, Law Commission No 154 (1986)).

This defence of Marxism carries over , however , to a critique of Modernism itself: not merely for the fact that Nietzsche attaches 'importance' to art but for the further claim ascribed to him, and therefore to Modernism itself , that 'the nature of aesthetic experience contains in nuce the form of understanding proper to the world itself (p. 66).

Pattern/rule:

Block C

Of course , these pathological tendencies are still with us today and are widely recognized in individual mental illness; but what is not generally appreciated is their extent in cultural forms, not all of which are limited to ancient societies.

Fig. 22: Sample worksheet with scaffolding

To create your own worksheet with scaffolding, we recommend copying the individual concordance lines into a Word document. You can do so by clicking on the document icon at the end of the concordance line to copy it (Fig. 23) and then paste it into your document (e.g., by using the keyboard shortcut ctrl + V).

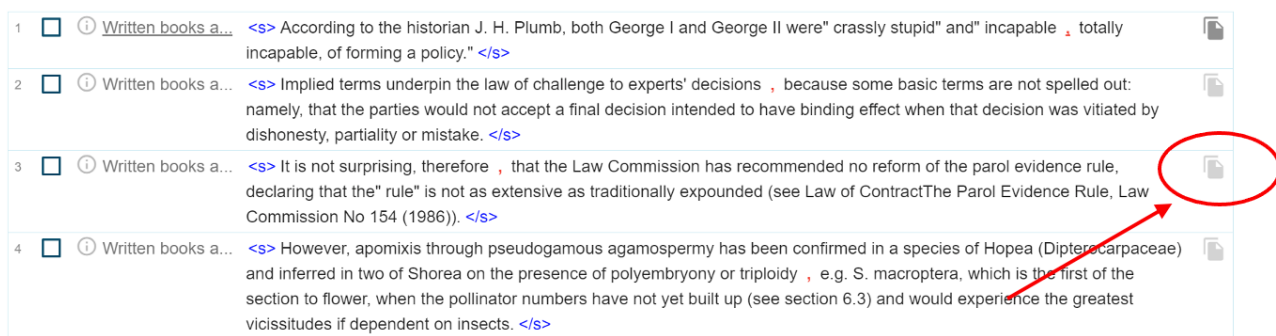


Fig. 23: Copying individual concordance lines

3.2 Lesson plan

Table 1: Lesson plan

Phase	Task	Interaction form	Materials/ Medium
Introduction: 5 min	Talking about the significance of commas; introduction to how a corpus can help with comma placement, e.g., in academic texts	Silent impulse (Plenum)	Picture "Punctuation", smartboard
Task 1: 25 min	Identifying patterns in comma placement in academic writing by analysing concordance lines Presentation and discussion of results Recourse/recap: explaining the need for commas and the deduced comma rules with reference to the picture from the beginning	Bus stop (1st individual work, 2nd partner work) Plenum	Worksheet Task 1 Smartboard Picture "Punctuation"
Task 2: 30 min	Applying the deduced rules to an academic text which lacks commas <i>(Optional end of lesson)</i>	Think-Pair-Share (1st individual work, 2nd partner work, 3rd Plenum)	Worksheet Task 2, notes from task 1
Task 3: 15 min	Students correct their own argumentative essays, focusing on comma placement	Individual work (Upcoming lesson: partner feedback)	Notes from task 1

3.2.1 Tasks

Introduction/Silent impulse:



Fig. 24: Silent impulse without commas



Fig. 25: Silent impulse with commas

Task 1:

In the worksheet for task 1 (cf. Fig. 21/22, Appendix), students are asked to examine concordance lines and identify patterns in order to derive rules for the use of commas in English. From these concordances, the following rules for comma placement may be deduced:

When to use commas – rules:

1. A comma is used after the introductory conjunctive adverbs (e.g., *however, moreover, furthermore*)
2. When these conjunctive adverbs are used in the middle of a sentence (i.e., as parenthetical

elements), two commas need to be placed: one before and one after the adverb.

3. A comma is used after introductory elements or phrases (e.g., *of course*, *in that case*, *on the other hand*).
4. Commas are used in enumerations. They separate items in a list consisting of more than two items.
5. A comma is used in relative clauses before the relative pronouns *which* or *who* when the pronoun introduces a non-essential element.
6. A comma is used before coordinating conjunctions joining two independent clauses (e.g., *but*, *as*, *yet*).
7. Two commas are used to separate insertions or parenthetical remarks, which add extra information and can be removed without changing the meaning of the sentence.
8. In conditional sentences starting with an *if*-clause, a comma is used to separate the *if*-clause from the main clause.

Task 2:

In worksheet 2, students apply the comma rules they formulated in task 1 to an argumentative essay by adding the missing commas in the right places. The essay has been adapted from the Michigan Corpus of Upper- Level Students Papers (MICUSP 2009).

Commas in Argumentative Writing II	
<p>Task 2:</p> <p>Read the excerpt from the argumentative essay “The Movement of Homosexual Marriage into the Public Sphere” (MICUSP 2009). All commas have been removed.</p> <ol style="list-style-type: none">Put in commas, following the comma rules which you discovered before.Name the rule for comma placement for each comma/set of commas you add to the essay.Choose three commas and explain their purpose and effect on the phrase in which they occur.	<p>The Movement of Homosexual Marriage into the Public Sphere</p> <p>“Marriage is and always will be a private contract between individuals. The boundaries of marriage in our country have been concrete for centuries. However they are currently being questioned. Should marriage remain defined as the union between a man and a woman? ... As Horowitz (2004) discusses there is a difference between the private sphere and public sphere in a relationship and according to him homosexuals are perfectly acceptable if they limit themselves to the private sphere ...</p> <p>In an expressive dimension the roles of homosexuals as a part of society’s norm are still being defined ... Legally homosexual couples have the right to be treated equally under the law because as I will argue their liberty is otherwise lost ... To neglect their right and the benefits of marriage is taking away the acknowledgement of their partnership which is valued in our society ...</p> <p>Rights as I am speaking of them are the certain benefits that have already been outlined in each state’s marriage license and grounded in each state’s legislature that our government is obligated to give each individual when he or she becomes married. These rights include but are not limited to: reductions in housing costs income taxes pensions and insurance premiums (Horowitz, 2004, p.465). The right to adopt a child is ... grant[ed] to heterosexual couples but is difficult for homosexual couples ... The importance of adoption by homosexual couples is an important topic but will not be of discussion in this argument ...</p> <p>As marriage becomes more culturally oriented the social context of marriage keeps changing. This reflects a change in many people’s belief systems but these beliefs can never change the rights or opportunities of any individual. These rights include but are not limited to the right to choose your own partner regardless of sex and society is still conforming to this realization. Also if Horowitz allows homosexuals to behave however they please in the privacy of their homes he is recognizing that there is nothing fundamentally wrong with the two as a couple he just does not wish to see them ...</p> <p>The norm that homosexual couples would be incapable parents for example due to social stigmas is the fault of society not their own. Recognition of homosexual marriage as a norm is the first step to removing these stigmas and others from society.”</p>

Fig. 26: Sample worksheet for the second task

Task 3:

Here, students are asked to check their own argumentative essays for comma placement. This helps them to internalise the patterns discovered in the previous tasks and shows them how they can make practical use of these rules to improve their own writing.

Task 4:

Note: This task is optional and not included in the lesson plan. However, it might be useful for students who tend to use more commas than necessary.

Commas in Argumentative Writing III

Task 4:

Read the excerpt from the argumentative essay “Stage Directions” written by a **native German English L2 learner**. This excerpt includes six extra commas that are not needed according to the rules for comma placement discovered in task 1.

Highlight/underline the extra commas.

Argumentative essay: Stage directions

Looking at famous plays, such as “Death of a Salesman” by Arthur Miller, it becomes clear, that the playwright included many stage directions in his play. Now the question is, if these stage directions are important for the readers or, if they should be left out.

Moreover, stage directions are often very specific, for instance, the Latin word “exeunt” is used, if more actors or actresses leave the stage. That is why stage directions are hard to understand and might confuse the reader.

However, there are also advocates, who praise the fact, that due to stage directions the reader gets to know what the characters, their gesture, and mimic look like.

Fig. 27: Sample worksheet for the fourth task

3.2.2 Methodological thoughts

In the introductory phase of the lesson, the silent impulse aims to encourage students to think of commas as a semantic, phrasal, and structural necessity, as they discover how commas can entail a change of sentential meaning (as in “I like cooking my grandma and my English teacher”).

It is the ‘form-focused approach’ which inspired the tasks (Ellis 2006: 94, Spada & Lightbown 2008: 185). Via the ‘Think-Pair-Share’ method, students continuously receive feedback, not only from the teacher but also from their classmates (Eisenmann 2019: 92-96), which represents a vital aspect of form-focused instruction (Ellis 2006: 94). Besides, they practise giving feedback to others, which – alongside the collaborative work – trains the learners’ reflective, evaluative, and communicative competences (Eisenmann 2019: 92). In

addition, the 'Bus stop' method considers students' individual pace of learning, matching those students with a similar pace and similar needs for the consequent partner work (Eisenmann 2019: 96-97).

In the framework of a planned focus on form (Ellis 2006: 100), students purposefully elicit frequent comma placements from the concordance lines they are presented with (Ellis 2006: 100). This bottom-up learner-corpus interaction gives rise to a natural, inductive, and experience-based second language acquisition (Spada & Lightbown 2008: 182-183, Tomlinson 2013: 14). Hence, students' initiative and commitment is promoted in the tasks (register: written academic).

Additionally, Task 3 enables students to improve their own argumentative essay as a final product, which represents the action-oriented lesson approach. Learners are more likely to acquire sustainable knowledge based upon their own insights rather than on prefabricated input.

Lastly, by choosing an argumentative essay on "The Movement of Homosexual Marriage into the Public Sphere" (MICUSP 2009), this lesson plan deliberately positions itself against the 'PARSNIP' approach (frequently applied by English textbooks), which promotes the avoidance of 'controversial' topics, such as politics, alcohol, religion, sex and sexual orientation, narcotics, -isms (e.g., communism, racism, etc.), and pork (hence the acronym 'PARSNIP'; Harwood 2010: 11, 434). This is because English language teaching should address all topics relevant to students' lives and society in general, enabling them to discuss these competently and express their own opinion. Hence, this lesson contributes to students becoming proficient English language speakers and autonomous language learners who are able to reflect on and talk about controversial topics.

3.2.3 Follow-up activities

As the proposed lesson plan ends with students writing their own essays, a constructive and beneficial follow-up activity would be a peer review of these essays. The students should be asked to bring their written essay to the following lesson, in which the teacher collects them and then distributes them among the students. During the peer reviewing process, the students read through the essays of their peers, checking for grammatical errors and especially paying attention to comma placement. Not only does this improve the individual essays, but it also helps students internalise and apply the rules for comma placement. Additionally, this strengthens students' proofreading competence and, thus, their ability to check written texts, identify errors, and correct them. However, it has to be taken into account that students will probably find some errors, but not all; they may even add more errors. This means that the essays may need to be reviewed by the teacher as well.

Another useful follow-up activity that deals with the problem of overusing commas is task 4 (cf. 3.2.1). This task consists of an authentic text written by a German student learning English that has too many commas. The superfluous commas should be identified and highlighted by the students.

In addition, follow-up lessons could also address the content of the argumentative essay, i.e., homosexual marriage. This could be done by having the students collect the arguments of the essay and adding their own as well as researching the legal situation in their home country.

4 Options and further ideas

The presented lesson plan can be supplemented with activities focusing on how to use corpora. This would constitute a holistic approach, which not only uses corpora as a means to an end, but also makes corpus tools and methods a central topic. Including activities that explain and show students how the data featured on the worksheet was compiled and letting them interact with the corpus directly provides them with

valuable study skills and techniques. In this way, they will be able to use corpora independently to check their own language use whenever they need to.

Furthermore, the presented lesson idea may be applied to other punctuation issues, such as the use of semicolons, for instance. Instead of looking at patterns on the word level, here, the focus would be on the sentence structure and the clause level of the concordance lines.

Moreover, this approach may be extended and used to contrast different registers. This could be done by comparing the use of commas in the “Written Academic” subcorpus (cf. 3.1.1) with comma use in a different register. Hence, differences in comma placement between academic texts and periodicals, for instance, may be analysed.

Apart from registers, this approach can also be used to contrast corpora in general. For students or teachers interested in differences in comma placement in different the varieties of English, it might be interesting to carry out a comparison of comma placement in the COCA, which represents American English, and the BNC. Another idea would be to contrast the patterns with those of a learner corpus. This can raise teachers’ awareness of frequent learner errors in this area and enable them to act accordingly.

Lastly, the approach taken in this chapter may be applied to any foreign language classroom. Provided that there is an available corpus for the language in question, this lesson is equally applicable to other languages that make use of commas, such as French or German.

5 Caveats and limitations

The lesson involves certain restrictions for both teachers and students. Teachers should be aware that, unless they are provided access via their school, Sketch Engine is limited to a 30-day trial period.

Therefore, it is necessary to plan well in advance in order to design further corpus-informed language teaching materials, e.g., by making use of the suggested ideas in this book.

A further potential problem concerns the selection of the concordance lines. Although selecting a random sample provides you with authentic language material, it is possible that some frequent uses of commas do not occur in the sample. Consequently, teachers must be particularly attentive when picking out the concordance lines. The selection should correspond to the frequency lists (Fig. 18, 20) but also suit the group of learners and their needs.

Finally, it should be mentioned that there are many other frequent uses of commas besides those discussed in this chapter.

6 Conclusion

We hope that the proposed chapter will be of help to teachers who want to design their lessons in a corpus-informed way in order to help students acquire ‘correct’ comma placement in academic texts. The lesson can be planned by teachers who have not had any experience with corpus interaction yet. Students need not have any experience either: the concordance lines will introduce students to corpora, familiarising them with corpus interaction. Following a bottom-up approach, the students derive frequent patterns of comma placement independently. This ensures that students engage intensively with the topic. Furthermore, students can apply their newly gained knowledge in tasks 2 and 3. Another advantage is the possibility of scaffolding in task 1 and the option to choose a different topic which corresponds more to the students’ interests in task 2. The proposed lesson is therefore both student-oriented and tremendously versatile.

7 Resources and references

The British National Corpus (BNC). Bodleian Libraries, University of Oxford, on behalf of the BNC Consortium. URL: <http://www.natcorp.ox.ac.uk/>. Available on sketchengine.eu.

Eisenmann, Maria. 2019. Teaching English: Differentiation and individualisation. Paderborn: Ferdinand Schöningh. <https://elibrary.utb.de/doi/book/10.36198/9783838551098> (14 March, 2021).

Ellis, Rod. 2006. Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly* 40(1), 83-107. doi:10.2307/40264512. (14 March, 2021).

Harwood, Nigel (ed.). 2010. English language teaching materials: Theory and practice. New York: Cambridge University Press.

Michigan Corpus of Upper-level Student Papers (MICUSP). 2009. Ann Arbor, MI: The Regents of the University of Michigan. <https://micusp.elicorpora.info/search/view/?pid=PHI.G0.12.1>.

Niedersächsisches Kultusministerium. 2015. Kerncurriculum für das Gymnasium, Schuljahrgänge 5-10: Englisch. https://cuvo.nibis.de/cuvo.php?skey_lev0_0=Fach&svalue_lev0_0=Englisch&skey_lev0_1=Dokumentenart&svalue_lev0_1=Kerncurriculum&skey_lev0_2=Schulbereich&svalue_lev0_2=Sek+I&skey_lev0_3=Schulform&svalue_lev0_3=alle&docid=1197&p=detail_view (02 June, 2021).

Spada, Nina & Patsy M. Lightbown. 2008. Form-focused instruction: Isolated or integrated? *TESOL Quarterly*, 42(2), 181-207. <http://www.jstor.org/stable/40264447>. (14 March, 2021).

Tomlinson, Brian. 2013. Second language acquisition and materials development. In Brian Tomlinson (ed.), *Applied linguistics and materials development*. 11-30. London: Bloomsbury Academic. <http://dx.doi.org/10.5040/9781472541567.ch-002>. (14 March, 2021).

8 Appendix

Example Worksheets

Commas in Argumentative Writing I

In the following table, you can see sentences containing one or more commas. The sentences, or so-called concordance lines, have been collected using the British National Corpus. This means that these sentences stem from a range of different academic texts that present authentic language use.

Tasks:

1. a) Examine the concordance lines and identify patterns of comma placement in academic writing. In each sentence, the comma in question is highlighted.
- b) Deduce comma rules from the sentences.

(Hint: Are there specific words or constructions that require a comma? Where exactly is the comma placed?)

Details	Sentence
1	Written books a... <p> It is not surprising, therefore, that the Law Commission has recommended no reform of the parole evidence rule, declaring that the "rule" is not as extensive as traditionally expounded (see Law of Contract/The Parole Evidence Rule, Law Commission No 154 (1986)). </p>
2	Written books a... <p> Of course, these pathological tendencies are still with us today and are widely recognized in individual mental illness, but what is not generally appreciated is their extent in cultural forms, not all of which are limited to ancient societies. </p>
3	Written books a... <p> Alternatively, as at Old Carlisle, the simple tile magister sufficed, as for example on the inscription dedicated to Mercury by Amilius, magister vici Bodati (Vic-sur-Seille) in the territory of the Madonnetiennes. </p>
4	Written books a... <p> On the other hand, sketches of Creole embedded in a London English turn often correspond to the most salient parts of the utterance, those which the other parties to the conversation respond to. </p>
5	Written books a... <p> Remember that in the production of texts for non-examined purposes (e.g. committee reports, novels, paintings, films), rough versions, sketches and drafts are the norm. </p>
6	Written books a... <p> Scientific discoveries and inventions, art and music, and so on can also be "nested" on the time-line. </p>
7	Written books a... <p> Having run through the preventive health catechism (more exercise, less tobacco and alcohol, and better diet) the 1979 report bluntly dismissed the idea of replacing Britain's tax-funded health service with some form of insurance scheme. </p>
8	Written books a... <p> In the words of Sallis again, "If we know that the head's motive is to protect the school, and especially the more needy in it, from questions which are ill-informed, destructive or selfish, we have a clear objective which is to make sure we don't deserve or unwittingly attract such accusation (1988:17). </p>
9	Written books a... <p> This indicates an acceptance of the state structures, or a recognition of a need for the separation of Northern Ireland from the remaining state structure of Ireland, but at the same time a continuing opposition to protestant - loyalist dominance. </p>
10	Written books a... <p> Countless examples can be studied in living creatures: seaboard harbours, great castles and minor houses, ports, botches, soccer goals and pinball machines. </p>
11	Written books a... <p> The original position seems to presuppose not just a neutral theory of the good, but a liberal, individualistic conception according to which the best that can be wished for someone is the unimpeded pursuit of his own path, provided it does not interfere with the rights of others. </p>
12	Written books a... <p> Otherwise they were on their own, playing, but not playing a game for there was no structured game to play. </p>
13	Written books a... <p> This defence of Marxism carries over, however, to a critique of Modernism itself, not merely for the fact that Nietzsche attaches 'importance' to art but for the further claim ascribed to him, and therefore to Modernism itself, that 'the nature of aesthetic experience contains in nuce the form of understanding proper to the world itself' (p. 66). </p>

14	Written books a... <p> So things such as ourselves, which perceive and have conscious awareness, do so because they are (or have) minds. </p>
15	Written books a... <p> However, this leaves open the possibility that the Earth, Venus and Mars acquired most of their volatiles from volatile-rich bodies by post-accretionary venturing. </p>
16	Written books a... <p> Moreover, this possibility is strengthened by the neglect of surface repositories, because regardless of how the various volatiles were acquired it is likely that a smaller fraction of the inert gases are retained in such reservoirs than are carbon and nitrogen. </p>
17	Written books a... <p> Certainly, it is important to study bureaucracies as institutions in their own right, even if we would not go as far as the poet Alexander Pope who wrote: </p>
18	Written books a... <p> Richards holds that in poetry the function of "feeling" tends to dominate that of "sense", while Jakobson identifies a special "poetic" function, which can be found in many uses of language, but which dominates over other functions in poetry. </p>
19	Written books a... <p> This, the court said, was an irrelevant consideration which rendered his decision unlawful. </p>
20	Written books a... <p> The gap between them, according to Althusser, is mediated by art. </p>
21	Written books a... <p> However there are a number of important controlling factors. Firstly, the subject fixates centrally (usually he has to report a digit which is presented in order to confirm this central fixation); secondly, the stimulus is exposed for less than 200 milliseconds to avoid eye movement; thirdly, the visual angle of the stimuli from the fixation point should be between 2.5' and 5' (otherwise, confounding effects occur). </p>
22	Written books a... <p> Furthermore, Rossi and Rosadini (1967), who also used the Amytal technique, reported that only one patient out of 74 right handers, all of whom received injections on both sides, did not have exclusively left hemisphere speech dominance. </p>
23	Written books a... <p> In that case, the opportunity to earn remuneration arose when the appellant obtained the grant of a tenancy of a public house owned by and tied to a large brewery company. </p>
24	Written books a... <p> Now if the prediction that follows from this maze of premises turns out to be false (in our example, if the planet does not appear at the predicted location), then all that the logic of the situation permits us to conclude is that at least one of the premises must be false. </p>
25	Written books a... <p> If there are exclusively private events, in the sense that they are in principle, and not just empirically, inaccessible to more than one observer, it might be argued that such events could not be intelligibly claimed, let alone shown to be, subject to any laws, and this means that no rational explanatory model could be constructed for them. </p>

Use this part of the worksheet to formulate rules:

Worksheet Task 1

Commas in Argumentative Writing I (with scaffolding)

In the following table blocks, you can see sentences containing one or more commas. The sentences, or so-called concordance lines, have been collected using British National Corpus. This means that these sentences stem from a range of different academic texts that present authentic language use.

Tasks:

1. a) Examine the concordance lines and identify patterns of comma placement in academic writing. In each sentence, the comma in question is highlighted. You can find one pattern in the sentences in each block.
- b) Deduce comma rules from the sentences in each block. (Hint: Are there specific words or constructions that require a comma? Where exactly is the comma placed?)

Block A

However, this leaves open the possibility that the Earth, Venus and Mars acquired most of their volatiles from volatile-rich bodies by post-accretionary venturing.

Moreover, this possibility is strengthened by the neglect of surface repositories, because regardless of how the various volatiles were acquired it is likely that a smaller fraction of the inert gases are retained in such reservoirs than are carbon and nitrogen.

Furthermore, Rossi and Rosadini (1967), who also used the Amytal technique, reported that only one patient out of 74 right handers, all of whom received injections on both sides, did not have exclusively left hemisphere speech dominance.

Pattern/rule:

Block B

It is not surprising, therefore, that the Law Commission has recommended no reform of the parole evidence rule, declaring that the "rule" is not as extensive as traditionally expounded (see Law of Contract/The Parole Evidence Rule, Law Commission No 154 (1986)).

This defence of Marxism carries over, however, to a critique of Modernism itself, not merely for the fact that Nietzsche attaches 'importance' to art but for the further claim ascribed to him, and therefore to Modernism itself, that 'the nature of aesthetic experience contains in nuce the form of understanding proper to the world itself' (p. 66).

Pattern/rule:

Block C

Of course, these pathological tendencies are still with us today and are widely recognized in individual mental illness, but what is not generally appreciated is their extent in cultural forms, not all of which are limited to ancient societies.

On the other hand, sketches of Creole embedded in a London English turn often correspond to the most salient parts of the utterance, those which the other parties to the conversation respond to.

In that case, the opportunity to earn remuneration arose when the appellant obtained the grant of a tenancy of a public house owned by and tied to a large brewery company.

Pattern/rule:

Block D

Remember that in the production of texts for non-examined purposes (e.g. committee reports, novels, paintings, films), rough versions, sketches and drafts are the norm.

Scientific discoveries and inventions, art and music, and so on can also be "nested" on the time-line.

Countless examples can be studied in living creatures: seaboard harbours, great castles and minor houses, ports, botches, soccer goals and pinball machines.

Pattern/rule:

Block E

In the words of Sallis again, "If we know that the head's motive is to protect the school, and especially the more needy in it, from questions which are ill-informed, destructive or selfish, we have a clear objective which is to make sure we don't deserve or unwittingly attract such accusation (1988:17).

So things such as ourselves, which perceive and have conscious awareness, do so because they are (or have) minds.

Richards holds that in poetry the function of "feeling" tends to dominate that of "sense", while Jakobson identifies a special "poetic" function, which can be found in many uses of language, but which dominates over other functions in poetry.

Pattern/rule:

Block F

This indicates an acceptance of the state structures, or a recognition of a need for the separation of Northern Ireland from the remaining state structure of Ireland, but at the same time a continuing opposition to protestant - loyalist dominance.

The original position seems to presuppose not just a neutral theory of the good, but a liberal, individualistic conception according to which the best that can be wished for someone is the unimpeded pursuit of his own path, provided it does not interfere with the rights of others.

This was evident not only in its mass organization, but also in the efforts to address working-class audiences and write them into the discourse of moral reform.

Pattern/rule:

Block G

This, the court said, was an irrelevant consideration which rendered his decision unlawful.

The gap between them, according to Althusser, is mediated by art.

Pattern/rule:

Block H

Now if the prediction that follows from this maze of premises turns out to be false (in our example, if the planet does not appear at the predicted location), then all that the logic of the situation permits us to conclude is that at least one of the premises must be false.

If there are exclusively private events, in the sense that they are in principle, and not just empirically, inaccessible to more than one observer, it might be argued that such events could not be intelligibly claimed, let alone shown to be, subject to any laws, and this means that no rational explanatory model could be constructed for them.

Pattern/rule:

Worksheet Task 1 (with scaffolding)

Task 2:

Read the excerpt from the argumentative essay "The Movement of Homosexual Marriage into the Public Sphere" (MICUSP 2009). All commas have been removed.

- a) **Put in** commas, following the comma rules which you discovered before.
- b) **Name** the rule for comma placement for each comma/set of commas you add to the essay.
- c) **Choose** three commas and **explain** their purpose and effect on the phrase in which they occur.

The Movement of Homosexual Marriage into the Public Sphere

"Marriage is and always will be a private contract between individuals. The boundaries of marriage in our country have been concrete for centuries. However they are currently being questioned. Should marriage remain defined as the union between a man and a woman? ...

As Horowitz (2004) discusses there is a difference between the private sphere and public sphere in a relationship and according to him homosexuals are perfectly acceptable if they limit themselves to the private sphere ...

In an expressive dimension the roles of homosexuals as a part of society's norm are still being defined ... Legally homosexual couples have the right to be treated equally under the law because as I will argue their liberty is otherwise lost ... To neglect their right and the benefits of marriage is taking away the acknowledgement of their partnership which is valued in our society ...

Rights as I am speaking of them are the certain benefits that have already been outlined in each state's marriage license and grounded in each state's legislature that our government is obligated to give each individual when he or she becomes married. These rights include but are not limited to: reductions in housing costs income taxes pensions and insurance premiums (Horowitz, 2004, p.465). The right to adopt a child is ... grant[ed] to heterosexual couples but is difficult for homosexual couples ... The importance of adoption by homosexual couples is an important topic but will not be of discussion in this argument ...

As marriage becomes more culturally oriented the social context of marriage keeps changing. This reflects a change in many people's belief systems but these beliefs can never change the rights or opportunities of any individual. These rights include but are not limited to the right to choose your own partner regardless of sex and society is still conforming to this realization.

Also if Horowitz allows homosexuals to behave however they please in the privacy of their homes he is recognizing that there is nothing fundamentally wrong with the two as a couple he just does not wish to see them ...

The norm that homosexual couples would be incapable parents for example due to social stigmas is the fault of society not their own. Recognition of homosexual marriage as a norm is the first step to removing these stigmas and others from society."

Worksheet Task 2

Worksheet Task 2: Solution

Task 4:

Read the excerpt from the argumentative essay “Stage Directions” written by a **native German English L2 learner**. This excerpt includes six extra commas that are not needed according to the rules for comma placement discovered in task 1.

Highlight/underline the extra commas.

Argumentative essay: Stage directions

Looking at famous plays, such as “Death of a Salesman” by Arthur Miller, it becomes clear, that the playwright included many stage directions in his play. Now the question is, if these stage directions are important for the readers or, if they should be left out.

Moreover, stage directions are often very specific, for instance, the Latin word “exeunt” is used, if more actors or actresses leave the stage. That is why stage directions are hard to understand and might confuse the reader.

However, there are also advocates, who praise the fact, that due to stage directions the reader gets to know what the characters, their gesture, and mimic look like.

Worksheet Task 4

Worksheet Task 4: Solution

1 Note: Sketch Engine randomises concordance lines with an algorithm that will always produce the same results if you enter the same number of concordance lines to be sampled. This ensures that the results are reproducible. Also, it ensures that students can follow along and obtain exactly the same results.

2 This corpus is freely accessible; use the following URL to get to the search interface:
<https://elicorpora.info/>.

3 Examples of French reference corpora available on Sketch Engine: the French Web (frTenTen) or the French Timestamped JSI web corpus.

4 Examples of German reference corpora available on Sketch Engine: the German Web (deTenTen) or the German Timestamped JSI web corpus.

PART IV

LESSON IDEAS: CONTENT AND LANGUAGE INTEGRATED LEARNING (CLIL) AT SECONDARY SCHOOL

Corpus-informed EFL lessons for Content and Integrated Learning (CLIL) at secondary school

- Teaching geography in English: How to describe a graph by **Hanna Bavendiek** and **Fabian Giesholt**
- “Black Lives Matter”: Collecting vocabulary to improve intercultural communicative competence by **Tilza Maria Meise-Reckefuß**
- Learning words and collocations independently: Rising temperatures and climate science by **Lena Janzen** and **Lisa Hörnschemeyer**
- Act responsible, think sustainable – Expanding vocabulary to write about sustainability by **Rieke Stahnke, Marlena Wegesin,** and **Lara Brede**
- Building a corpus to create materials for a debate on pollution by **Marie Gröninger** and **Ha Mi Nguyen**

9. Teaching geography in English: How to describe a graph

HANNA BAVENDIEK AND FABIAN GIESHOLT

1 Introduction and rationale

In a fast and globalised world, English serves as the lingua franca of the 21st century, especially on the internet. Bilingual teaching of other subjects besides English itself is finding increasing popularity among German secondary schools, as it offers great potential for successful second language acquisition.

Geography presents a perfect opportunity to combine regular class topics with students' increasing language proficiency. Due to the internationality of its content, intercultural learning can be particularly encouraged when geography is taught in English. Moreover, the vivid nature of geography teaching, which frequently involves lots of maps, illustrations and caricatures, provides suitable support for second language learners because geography is often not perceived to be as abstract as many other school subjects (Niedersächsisches Kultusministerium 2015: 28). The contemporary relevance of geographic topics presents promising opportunities for bilingual teaching. The English curriculum of Lower Saxony (Germany) further specifies that appropriate subject-specific terminology ought to be acquired in both the first and second language of learners (Niedersächsisches Kultusministerium 2015: 28).

A central aspect of geography teaching in the classroom is the analysis of graphics and illustrations. Especially graphs and diagrams are found in almost all geography topics. Thus, the ability to describe such figures is a crucial competence in order to succeed in geography. However, in a foreign language context, this ability can be quite difficult to achieve. Students may be able to produce precise descriptions in their mother tongue but struggle to transfer this ability to English because specific vocabulary is missing. In light of this difficulty, a corpus analysis can help retrieve the language necessary to describe graphs and diagrams in an idiomatic manner.

A corpus analysis of geography-specific language offers a wide variety of options for students to improve their own language skills. Following a data-driven learning approach, students are enabled to research collocations and frequent patterns of native and expert language usage themselves. Furthermore, these methods of corpus analysis are easily deployable within other language learning processes to increase language proficiency in other areas of English communication. Corpus analysis enables students to gain insights into native/expert language usage through independent, self-regulated learning.

Outline and objectives

This lesson is designed to be taught within a bilingual geography classroom of 10th-graders at a German secondary school (Gymnasium). Bilingual teaching combines foreign language acquisition

with specific subjects to acquire a more profound language proficiency within a subfield of academic study. In this lesson, students learn about the demographic transition model by reading texts and drawing a graph. Furthermore, they learn how to describe a graph by querying a corpus to advance their language skills.

By the end of the lesson, students will be able to ...

- ... understand and use appropriate words and collocations to describe a graph.
- ... use the Concordance, Thesaurus, and Word Sketch function in Sketch Engine.
- ... differentiate between these functions and know when to use them.

The student activity in relation to diagrams is split into two 90-minute lessons. First, the diagrams are created from the necessary information in worksheet 1 (see page 10). Second, the actual connection to a corpus-based analysis is drawn: students analyse concordances in the English Web 2015 (enTenTen15) corpus on Sketch Engine. This lesson takes a learner-corpus interaction approach towards gaining deeper insight into language usage by utilizing corpora for specific contexts.

2 Corpus, tools and methods

The planned lesson is based on the online corpus tool Sketch Engine. Sketch Engine provides a variety of different corpora and functions for corpus analysis, making it a suitable option for language learners to engage with. The students in this particular class are already familiar with the online tool. The following step-by-step-instructions may provide further support. For the first part of the corpus-based lesson, students work with the English Web 2015 (enTenTen15) corpus and use the Concordance and Thesaurus functions to search for synonyms. In the second part of the lesson, students use the Word Sketch function to find topic-related collocations, which they can use to describe the demographic transition model. Therefore, students will work with three different functions of Sketch Engine, which will help them to increase their language proficiency in the bilingual geography classroom. Utilizing the different functions and corpora will show students that one needs to select the right function to obtain meaningful results.

3 Step-by-step guide for Teachers

3.1 Lesson Plan

Preparatory lesson

Table 1: Lesson Plan ‘The demographic transition model’¹

	Content	Methods	Materials
Warm-up	Photo description	Plenary discussion	Photo of babies, projection via a projector
Task 1	Students read the texts and transform the information into a diagram	Group work	Worksheets
Task 2	Students draw the diagram on the board and explain the stages of the demographic transition model	Plenary discussion	Board
Homework	Underline all the words in the texts that describe the demographic transition model		

Main lesson

Table 2: Lesson Plan: “How to describe a graph”

	Content	Methods	Materials
Warm-up	Description of the demographic transition model Teacher asks pupils to use other verbs than <i>show</i> , teacher makes notes of them on the board	Plenary discussion	Demographic transition model, projection via a projector, board
Task 1	Students work with corpora to find useful words for describing the diagram	Pair work	Worksheets, Computers, Internet
Task 2	Students use Sketch Engine to search for alternative terms to <i>fall</i> and <i>increase</i> to write a description of the patterns of population development	Individual work	Worksheet, Computers, Internet
Homework	Write the description		

3.2 Instructions

In the preparatory lesson, students learn about the characteristics of the different stages of the demographic transition model, following the information provided on Worksheet 1 to draw a graph. This provides the basis for the following lesson with regard to content. As a preparation for the main corpus-based lesson, students get homework in which they have to underline all the words in the texts on Worksheet 1 that describe the demographic transition model. The idea of the corpus-based lesson is essentially to help students identify different words and collocations to describe a graph by using a corpus.

3.3 Main Lesson

3.3.1 Warm-up

The lesson begins with an oral description of the demographic transition model from the previous lesson in a plenary discussion. We expect students to use the word *show* most frequently, e.g., *the graph shows...* The teacher will point this out and then students are encouraged to think of verbs other than *show* that can be used in this context. Meanwhile, the teacher notes down the students' ideas on the board.

3.3.2 Main Lesson Task 1

The following tasks require access to web-enabled computers or tablets in the classroom. Students work in pairs and share a device. They follow the instructions on Worksheet 2 and work independently with Sketch Engine. Students will have previously received a general introduction to Sketch Engine and created a personal account. However, the worksheet still provides detailed instructions and extra tips.

In the first task, students use the Concordance function in Sketch Engine to search for concordance lines that include the phrase: *graph shows*. By looking at the type of diagram that is described and identifying its source of origin, students familiarize themselves with working with a corpus. Seeing the structure of the concordance lines will help them to understand that a corpus is composed of various texts from different sources.

Next, students search for synonyms of the verb *show* in the English Web 2015 (enTenTen15) corpus by using the Thesaurus function. They note down their findings and draw assumptions about which verbs may be used in the context of graph descriptions.

To verify or disprove students' assumptions, the next step includes the use of another function of Sketch Engine. This time students do not use the Thesaurus function but rather the Word Sketch function to look for collocations by searching for verbs that are used with *graph* as a subject. Then students write down their results and compare them to their findings from the Thesaurus search. Additionally, they may reflect on the differences between a Thesaurus and a Word Sketch search and decide which method is more effective for their purpose of finding alternative verbs to *show* to describe a graph.

3.3.3 Main Lesson Task 2

The final task of this corpus-based lesson attempts to encourage students to apply their knowledge of working with a corpus to write in-depth descriptions of graphs and their trends and peculiarities. For homework, students were asked to underline all the words that describe the graph on worksheet 1. Apart from *show*, the verbs *increase* and *fall* are also used to describe the developments of the stages of the demographic transition model on the worksheet. Students are asked to use Sketch Engine to search for synonyms of these verbs and to find additional words and collocations that describe the graph. Finally, students write a complete description of the demographic transition model.

3.4 Tasks

Worksheet 1

Task 1: Read the text from M1 and, based on the information from the text, draw lines illustrating the birth rate, death rate and population growth into M2.

M1 The five stages of demographic development

Stage 1: High stationary	
High death rate due to famines, diseases and a lack of medical care.	The birth rate is also high due to a lack of birth control, early marriage age for women and the necessity of child labour so that they can contribute to the family income.
The population increase is low as the natural increase is rather low. The UK was in stage 1 before 1760. Only a few places are in stage 1 nowadays – maybe a few remote tribes in the tropical rainforests.	

Stage 2: Early expanding	
The death rate is starting to fall due to improved medicine, cleaner water, more and better nutrition and better sanitary conditions.	The birth rate is still high – same reasons as above.
The natural increase is high – the population increases fast. The UK was in stage 2 from about 1760 to the end of the 18 th century, during the Industrial Revolution. Today, a number of Less Economically Developed Countries (LEDCs) are in this stage, e.g., Ethiopia and Cambodia.	

Stage 3: Late expanding	
The death rate is still falling, for the same reasons as above.	Now the birth rate is starting to fall as there are fewer farmers needing children to work on the family land; infant deaths are falling; birth control is available and women have the opportunity to receive education and are getting married later.
There is still some natural increase, but it is lower than it used to be. As a result, the population increase is slowing down. The UK was in stage 3 between 1900 and about 1950. Today, many LEDCs are in this stage, e.g., Brazil, Indonesia and Mexico.	

Stage 4: Low stationary	
The death rate remains low.	The birth rate is low due to birth control. People are now having the number of children they want to have.
There is little or no natural increase. Therefore, the population does not grow much. The UK has been in stage 4 since about 1950. Many More Economically Developed Countries (MEDCs) are momentarily in this stage, e.g., Japan, the USA, France. Only a few LEDCs have reached this stage.	

This worksheet was downloaded from: Bavendiek, Hanna & Fabian Giesholt (2021). "Teaching geography in English: How to describe a graph." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

Stage 5: Declining

The death rate could increase because a higher proportion of people is elderly.	The birth rate stays low and could become lower due to lifestyle changes. People have fewer children and do so later in life.
--	--

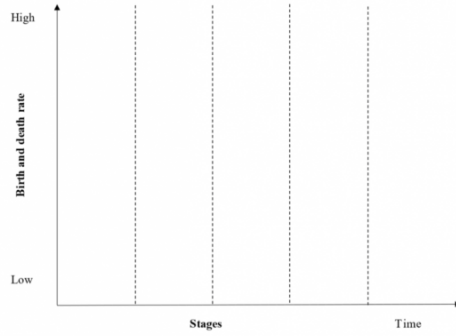
If more people die than are born, there is a negative natural increase so that the **population** falls.

This stage was not in the model when it was first created; it has been added to reflect current developments in population evolution.

The UK could reach this stage soon. Other MEDCs, such as Sweden, Germany and Italy, have entered this stage already.

(adapted from Niedersächsisches Kultusministerium 2014: 32)

M2 The demographic transition model



(adapted from Niedersächsisches Kultusministerium 2014: 33)

This worksheet was downloaded from: Bavendick, Hanna & Fabian Giesholt (2021). "Teaching geography in English: How to describe a graph." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

Worksheet 2

For the following tasks, you should work in pairs. Follow the instructions.

1. Connect to www.sketchengine.eu and log in to your account.
2. First you need to use the English Web 2015 (enTenTen15) Corpus.

TIP 1 To open the English Web 2015, click on the left side bar and click on 'select corpus'. Now select 'advanced' (on the top) and write 'English Web 2015' into the search field and select the corpus.

TIP 2 To get a random sample you can click on the dice with the question mark on it above the concordance lines.

3. Click on the Concordance function and type *graph shows* into the simple search field. Now search for results and draw a random sample of 200 concordance lines.

TIP 3 If you click on the red words, you can read the full sentence plus the surrounding sentences to get more information about the phrase.

Look through the concordance lines and choose one example. Write down the concordance line including the source of origin ('Details') and explain what type of graph is being described.

4. Go back to the left-hand side menu and click on the Thesaurus function. Click on the top bar and select 'Advanced'. Type the word *show* and select 'verb' as the part of speech. Click on 'GO'. Enter your findings in the table below.

Number	Synonym	Frequency in the EnTenTen15 corpus
1		
2		
3		
4		
5		

This worksheet was downloaded from: Bavendick, Hanna & Fabian Giesholt (2021). "Teaching geography in English: How to describe a graph." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

Number	Synonym	Frequency in the EnTenTen15 corpus
6		
7		
8		
9		
10		

5. Which of these verbs do you think can also be used to describe a graph? Which factors influenced your decision?
6. Next, select the Word Sketch function from the left-hand side menu. Click on 'Advanced'. Type in *graph* and select 'noun' as part of speech. Click on 'GO'. Now look at the column 'verbs with "graph" as subject'. Write down your results in the table below. Click on the three dots to the right of a word and click on 'Concordance' to view examples in context.

Number	Word	Example sentence
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

This worksheet was downloaded from: Bavendick, Hanna & Fabian Giesholt (2021). "Teaching geography in English: How to describe a graph." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

7. Compare your results from steps four and six. What are the main differences between a Thesaurus and a Word Sketch search? Which one was more effective for your purpose and why?
8. Now that you have identified verbs to describe graphs, your next task is to find words and phrases to describe graphs in further detail, including its trends and characteristics. Describe the demographic transition model and the three graphs in detail. Use words and phrases that go beyond *increase* and *fall*. But first, find these words using Sketch Engine. If you are struggling to find results, go back and reconsider which functions of Sketch Engine can be used to generate concordance lines. Identify at least five distinct phrases which are used to describe the trends of graphs in further detail and write down your result in the table. Finally, write a complete description of the demographic transition model.

Corpus	Word	Example sentence

This worksheet was downloaded from: Bavendick, Hanna & Fabian Giesholt (2021). "Teaching geography in English: How to describe a graph." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

4 Options and further ideas

In order to achieve a suitable level of difficulty for all students, a certain degree of internal differentiation is necessary. While some students might find it comparatively easy to work with Sketch Engine individually, others might struggle and need more guidance. Thus, it is possible to make tasks easier or more challenging by leaving out or adding intermediate steps in the task description. The latter could be achieved through the addition of screenshots of the actual web pages.

Another possibility is the introduction to further materials for quicker students. These students could analyse an additional graph on population growth and change. Students could describe a graph showing the development of a population in a country of their choice after researching it in their textbook, atlas, or online. Figure 1 depicts an example of the German demographic transition model.

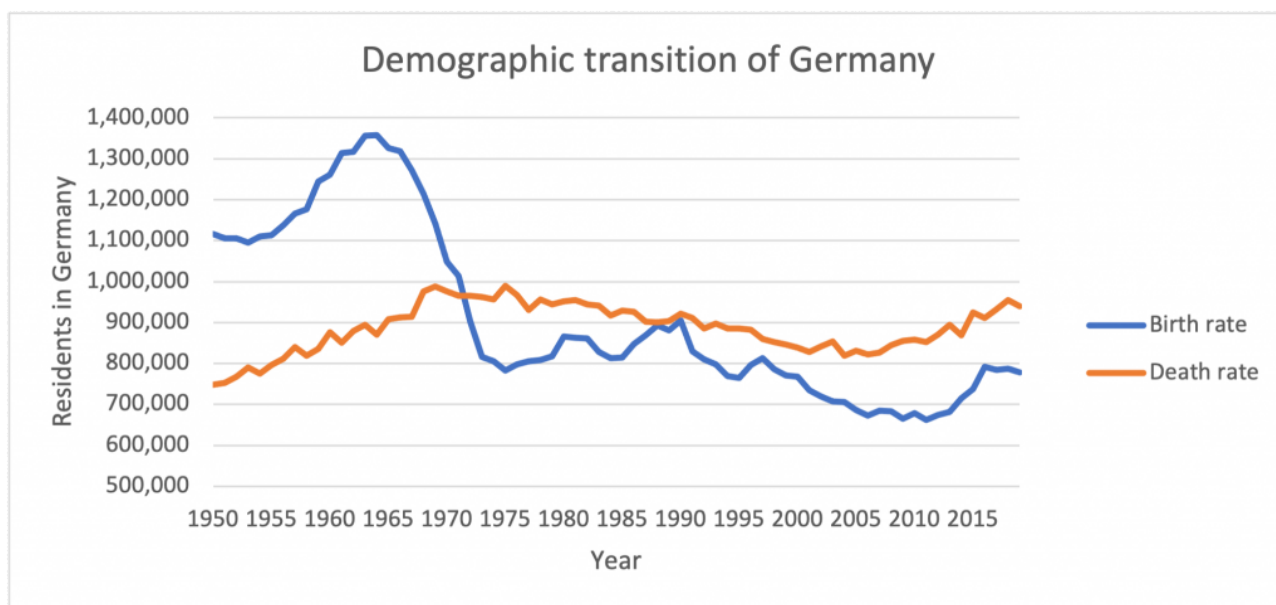


Fig. 1: Demographic transition model of Germany (own illustration based on data from Statistisches Bundesamt 2020)

Once the initial tasks are completed, students who have finished could do a similar corpus analysis to acquire further language knowledge to describe maps or caricatures. To enable them to elaborate on this, an exemplary map or caricature is needed. The teacher could provide these in advance and students could describe either one of them with the help of corpus searches.

5 Caveats and Limitations

Difficulties might occur when working with the commercial platform Sketch Engine if the teacher and the students do not have personal accounts. Even though Sketch Engine has an open version (app.sketchengine.eu/#open) as well, it is not sufficient for the requirements described in this chapter since it only allows access to a very limited selection of corpora. One option would be to make use of a free 30-day trial during which all of Sketch Engine's functions are accessible. However, this is only a short-term solution and might serve the purpose of this lesson but will restrict further corpus-learner interaction in the classroom. Alternatively, teachers may want to consider using the SKELL interface (skell.sketchengine.eu), which is much more limited in terms of functions, but is entirely free and was specifically designed with language learners in mind.

6 Conclusion

Bilingual teaching aims to encourage the most authentic use of a foreign language in a specific subject. We hope that this chapter has shown that working with corpora can be a great opportunity for students to expand their vocabulary and increase their language proficiency. Even though the corpus-based lesson itself requires more time than if the teacher simply handed out a list of vocabulary, we believe that corpus-based, data-driven learning is a great addition to the bilingual geography classroom. Searching words and collocations within a corpus challenges students. The lesson presented in this chapter thus requires students' cognitive engagement throughout the tasks. Besides, the use of digital media is encouraging for students and is likely to be a welcome change to the usual style of teaching that students experience on a daily basis. Students not only learn new words but also how phrases and collocations are used in a context that is based on real, authentic texts. Furthermore, learner-corpus interaction in the classroom equips students with a tool that provides opportunities for use in the future that they can benefit from to improve their language skills in the long term.

7 Resources and references

Campbell, Maxi-Ann. 2018. Using COCA to answer the question on everyone's lips. In Eric Friginal (ed.), *Corpus linguistics for English teachers: New tools, online resources, and classroom activities*, 299-305. New York, NY: Routledge.

Friginal, Eric. 2018. *Corpus linguistics for English teachers: New tools, online resources, and classroom activities*. New York, NY: Routledge.

Niedersächsisches Kultusministerium. 2014. Bilingualer Unterricht – Materialien für kompetenzorientierten Unterricht im Sekundarbereich I. http://arbeitsplattform.bildung.hessen.de/fach/bilingual/lehrer/ausb/mat/nibis-materialien_bili_2014.pdf (06 February, 2021).

Niedersächsisches Kultusministerium. 2015. Kerncurriculum für das Gymnasium, Schuljahrgänge 5-10: Erdkunde. https://cuvo.nibis.de/cuvo.php?skey_lev0_0=Schulbereich&svalue_lev0_0=Sek+I&skey_lev0_1=Fach&svalue_lev0_1=Erdkunde&docid=1112&p=detail_view (13 August, 2020).

Statistisches Bundesamt. 2020. Anzahl der Geburten und der Sterbefälle in Deutschland in den Jahren von 1950 bis 2019. <https://de.statista.com/statistik/daten/studie/161831/umfrage/gegenueberstellung-von-geburten-und-todesfaellen-in-deutschland/> (15 February, 2021).

Notes

1. Adapted from Niedersächsisches Kultusministerium 2014

10. "Black Lives Matter": Collecting vocabulary to improve intercultural communicative competence

TILZA MARIA MEISE-RECKEFUSS

1 Introduction and rationale

The death of George Floyd in the USA. has rekindled debates about racism and triggered protests around the world. George Floyd's last words, "I can't breathe", became a slogan to highlight the reality that non-white people have been facing in their everyday life, oppressed by racism and prejudice. Another slogan used around the world in the protests was "Black Lives Matter". The hashtag and movement Black Lives Matter was created in 2013 as a reaction to the acquittal of George Zimmerman, a neighborhood watch volunteer, who killed Trayvon Martin, an unarmed black teenager.

The local secondary school curriculum states that students should be able to recognize, question, and when necessary, relativize and revise their perceptions, judgments, and prejudices (Niedersächsisches Kultusministerium 2017:20).

To cover this aspect stated by the school curriculum, this chapter provides teachers/students with the necessary instructions to use corpus linguistic research for collecting vocabulary and collocations related to the "Black Lives Matter" movement. Students will therefore have the opportunity to learn vocabulary and collocations through contact with authentic written language on a topical issue.

The chapter gives the students some tools to talk about racism using respectful language. In this sense, the collected vocabulary is also an effective way of developing intercultural communicative competence and promoting anti-racism education. Sommer and Roiha (2017: 8) suggest that anti-racism education is best "conveyed through small steps and practices". They argue that the examination of language issues and of "what is said" can contribute to developing intercultural communication competence.

Outline and objectives

Target Learners: 10th-grade students

Duration: approx. 90 min

This chapter presents **two options** to use the corpus material in the lesson:

Option 1: Only the teacher engages with the corpus and designs teaching materials from the findings in the corpus using screenshots, extracted sentences, collocations, and patterns. You can find ready-made corpus-based teaching material in this chapter as well as instructions and ideas on

how to create your own material on “Black Lives Matter” or any other topic of your choice. In the lesson, students engage with the material to learn relevant language patterns and to develop a greater understanding of the topic.

Option 2: In this alternative version, the teacher shows students how to work with the corpus themselves. Students look for patterns and collocations in the corpus and discuss the results. They may also work with additional teaching materials prepared beforehand using the corpus. This approach is more hands-on and may provide students with a better insight into the topic and language patterns, but it is also more time consuming and requires students to have access to internet-enabled devices.

Learning Objectives:

- Students learn vocabulary and collocations related to the “Black Lives Matter” movement and to racism in our modern society more broadly.
- Students improve their ability to speak and write about the “Black Lives Matter” movement and to understand texts related to this topic.
- Students develop greater intercultural competence and increase their awareness of racism in our society.
- If you choose option 2: Students know how to query a corpus in order to find collocations on specialized, topical topic.

2 Corpus, tools and methods

According to Friginal (2018: 5), besides providing a contextualized use of language, the application of corpus linguistics in English language teaching allows students to perceive the variety of vocabulary and grammar found in authentic samples of written and spoken language: “Students will have to incrementally learn what linguistic variation in everyday language means, together with its sociocultural values”. Considering the topicality of the “Black Lives Matter” movement, it can be assumed that it will be almost impossible to find pedagogical materials on this topic. For that reason, the NOW Corpus was selected for this lesson. The NOW Corpus (News on the Web) contains 10.7 billion words from online news publications and new data is added on a daily basis.

This chapter will demonstrate how to create a virtual corpus from the data of the NOW corpus that consist of texts related to the “Black Lives Matter” movement and how teachers can exploit it to create their own materials on the basis of authentic texts and language patterns, and/or to conduct lessons in which students interact with the corpus themselves to find relevant language patterns.

3 Step-by-step guide: Teacher Preparation

3.1 Creating a virtual corpus

To collect vocabulary related to the “Black Lives Matter” movement, create a specialized virtual corpus for the “Black Lives Matter” movement using one of the two options explained below, or create two virtual corpora using both methods and then combining them (see end of section 3.1). Note that if you wish to let

your students engage with the specialized corpus to search for patterns, they will have to follow the same steps to create a virtual corpus on their own, as you cannot share corpora on english-corpora.org.

Note that the english-corpora.org interface regularly displays a notice encouraging you to upgrade your account if you are a non-paying user. But do not worry, the message will disappear after a few seconds and you can continue your search. A premium account is not necessary for this lesson.

Option 1:

1. Open your browser and go to <https://www.english-corpora.org/>.
2. Log in to your account or register by clicking on **my account**.
3. Click on **News on the Web (NOW)**.

The screenshot shows the English-Corpora.org website interface. At the top, there is a navigation bar with links for home, corpora, users, related resources, my account, upgrade, and help. Below the navigation bar, there is a section titled "The most widely used online corpora" with a list of corpora. The "News on the Web (NOW)" corpus is highlighted with a red box. Below this section, there is a table of corpora with columns for Corpus (online access), Download, # words, Dialect, Time period, and Genre(s).

Corpus (online access)	Download	# words	Dialect	Time period	Genre(s)
iWeb: The Intelligent Web-based Corpus	↓	14 billion	6 countries	2017	Web
News on the Web (NOW)	↓	12.8 billion+	20 countries	2010-yesterday	Web: News
Global Web-Based English (GloWbE)	↓	1.9 billion	20 countries	2012-13	Web (incl blogs)
Wikipedia Corpus	↓	1.9 billion	(Various)	2014	Wikipedia
Corpus of Contemporary American English (COCA)	↓	1.0 billion	American	1990-2019	Balanced
Coronavirus Corpus	↓	1066 million+	20 countries	Jan 2020-yesterday	Web: News
Corpus of Historical American English (COHA)	↓	475 million	American	1820-2019	Balanced
The TV Corpus	↓	325 million	6 countries	1950-2018	TV shows
The Movie Corpus	↓	200 million	6 countries	1930-2018	Movies
Corpus of American Soap Operas	↓	100 million	American	2001-2012	TV shows
Hansard Corpus		1.6 billion	British	1803-2005	Parliament
Early English Books Online		755 million	British	1470s-1690s	(Various)

Fig. 1: Accessing the NOW corpus

4. Click on **Texts/Virtual** and then on **Create corpus**.

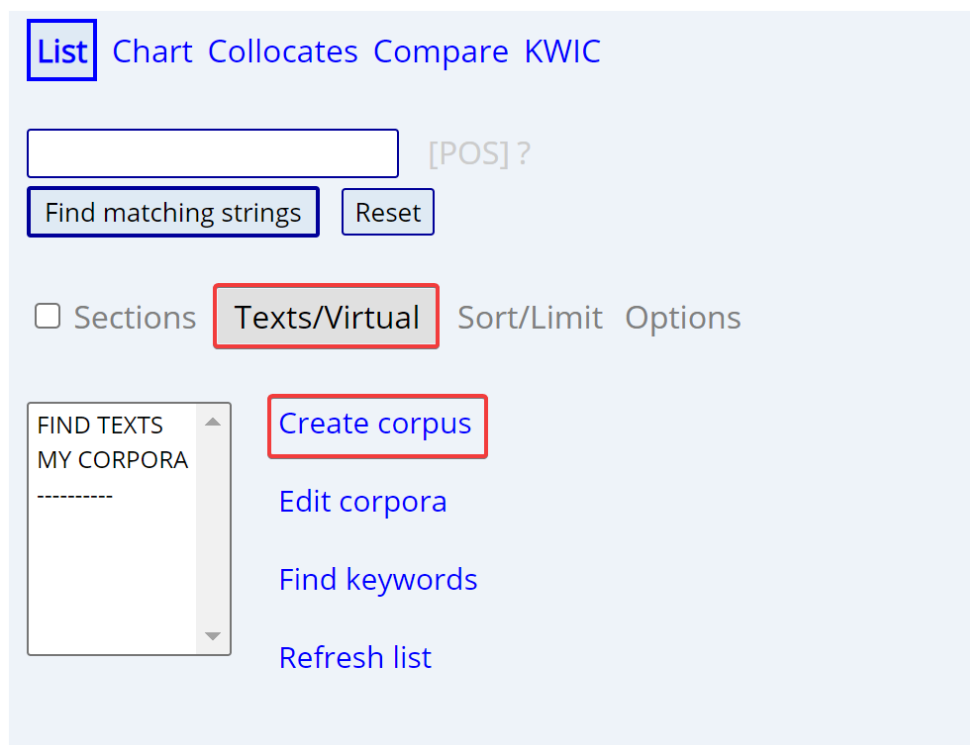


Fig. 2: Creating a virtual corpus

5. Now type the surname "Floyd" into the box next to **Article title**.
6. Select the **United States** as **Country**.
7. Select the date 05/25/2020 (which is the day of George Floyd's death) as the start date and today's date as the end date (in the screenshot: 07/08/2020). Note the use of the American date format: MM/DD/YYYY.
8. Now click on **Submit**.

Web domain	<input type="text"/>	<input type="button" value="Find sources"/>
(can use substring, e.g. Times, Houston)		
Article title	<input type="text" value="Floyd"/>	
Country	<input type="text" value="----"/> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> United States <input type="checkbox"/> Canada <input type="checkbox"/> Great Britain <input type="checkbox"/> Ireland 	
Dates	<input type="text" value="05/25/2020"/>	to <input type="text" value="07/08/2020"/>
Words in text	<input type="text"/>	
# texts (max)	<input type="text" value="1000"/>	
<input type="button" value="Submit"/> <input type="button" value="Reset"/>		

Fig. 3: Creating a corpus of articles about George Floyd

9. Deselect the texts that are not related to George Floyd's death (in my sample, I only found three: 1, 5, 14).
10. **Save as** "Black Lives Matter Floyd".
11. Click on **Submit**.

You can add or remove texts below, and then either save these texts as a new virtual corpus, or else add them to an existing virtual corpus.

SAVE AS: <input type="text" value="Black Lives Matter Floyd"/> OR ADD TO: <input type="text" value="--SELECT--"/>					
<input type="button" value="SUBMIT"/> <input type="button" value="RESET"/>					
HELP	<input type="checkbox"/> 100	DATE	COUNTRY	WEBSITE	TITLE
1	<input type="checkbox"/>	20-05-25	United States	New York Post	Floyd Mayweather partied at crowded Arizona nightc...
2	<input checked="" type="checkbox"/>	20-05-26	United States	Aclu.org	ACLU Responds to Minneapolis Police Killing George...
3	<input checked="" type="checkbox"/>	20-05-26	United States	Axios	Minneapolis unrest as hundreds protest death of Ge...
4	<input checked="" type="checkbox"/>	20-05-26	United States	Billboard	Ice Cube Wonders 'How Long' Before People 'Strike ...
5	<input type="checkbox"/>	20-05-26	United States	Billboard	Snoop Dogg, Justin Bieber, Ariana Grande & More Sp...
6	<input checked="" type="checkbox"/>	20-05-26	United States	Bleacher Report	Floyd Mayweather in AZ Club Amid COVID-19 'Disturb...
7	<input checked="" type="checkbox"/>	20-05-26	United States	Fox News	George Floyd's cousins speak out about heartbreak ...
8	<input checked="" type="checkbox"/>	20-05-26	United States	Patch	Medical Examiner: George Floyd Autopsy Results Are...
9	<input checked="" type="checkbox"/>	20-05-26	United States	rollingstone.com	Pink Floyd to Release Rarities on New 'Evolving Pl...
10	<input checked="" type="checkbox"/>	20-05-26	United States	The Boston Globe	Protests erupt in Minneapolis after George Floyd d...
11	<input checked="" type="checkbox"/>	20-05-26	United States	vulture.com	All 165 Pink Floyd Songs Ranked, From Worst to Bes...
12	<input checked="" type="checkbox"/>	20-05-26	United States	Wall Street Journal	Minneapolis Police Clash With Protesters Angered b...
13	<input checked="" type="checkbox"/>	20-05-26	United States	YAHOO!	Scottsdale mayor isn't happy with Floyd Mayweather...
14	<input checked="" type="checkbox"/>	20-05-26	United States	YAHOO!	Stephen Jackson emotional over friend George Floyd...
15	<input checked="" type="checkbox"/>	20-05-27	United States	abcnews.go.com	Sister of George Floyd, black man who died after p...
16	<input checked="" type="checkbox"/>	20-05-27	United States	al.com	DJ. Fluker reflects on the death of George Floyd ...
17	<input checked="" type="checkbox"/>	20-05-27	United States	buzzfeednews.com	George Floyd Death: Photos Capture Hundreds Of Peo...
18	<input checked="" type="checkbox"/>	20-05-27	United States	buzzfeednews.com	George Floyd Obituary: Friends And Family Describe...
19	<input checked="" type="checkbox"/>	20-05-27	United States	chicago.suntimes.com	Gov. I.B. Pritzker on Georse Floyd death: 'Being b...

Fig. 4: Submitting a search

12. Now you can see your corpus in **My Virtual Corpora**.
13. Refresh your browser to view your corpus in **Texts/Virtual**.

Option 2:

1. Go back to the search page.
2. Click on **Texts/Virtual**.
3. Type in "black lives matter".
4. Click on **Find texts**.
5. Submit your search by clicking on **Find matching strings**.

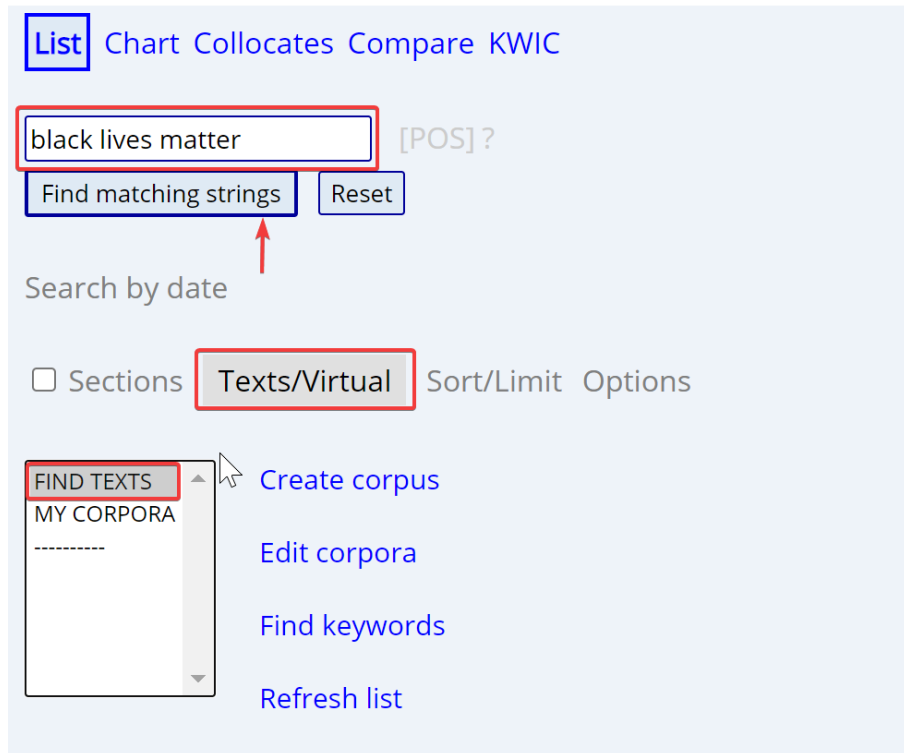


Fig. 5: Finding texts

6. On the next page, click on **Save List** and write "Black Lives Matter" into the box next to **SAVE AS**.
7. Click on **Submit**.

(Optional) SAVE AS: <input type="text" value="Black Lives Matter"/> OR ADD TO: <input type="text" value="--SELECT--"/> <input type="button" value="SUBMIT"/> <input type="button" value="RESET"/>					
HELP	<input type="checkbox"/>	<input type="text" value="100"/>	TEXT	# WORDS	# HIT
1	<input checked="" type="checkbox"/>		GB 20-12-16: INDEPENDENT.CO.UK	1738	8
2	<input checked="" type="checkbox"/>		US 20-06-05: TENNESSEAN	32877	6
3	<input checked="" type="checkbox"/>		US 20-05-30: POSTANDCOURIER.COM	4199	6
4	<input checked="" type="checkbox"/>		US 20-06-27: WUSA9	5500	5
5	<input checked="" type="checkbox"/>		US 20-08-15: LOS ANGELES DAILY NEWS	1924	4

Fig. 6: Saving the results as a list

- Now you can see your two corpora: one with articles about George Floyd and one with texts in which “Black Lives Matter” is mentioned.

MY VIRTUAL CORPORA [Detailed overview](#)

HELP			LIST NAME ↓	# TEXTS ↓	# WORDS ↓	FIND KEYWORDS <input checked="" type="radio"/> SPECIFIC <input type="radio"/> FREQ	CREATED ↓
1			BLACK LIVES MATTER	100	358,264	NOUN VERB ADJ ADV N+N ADJ+N	0 h
2			BLACK LIVES MATTER FLOYD	998	863,756	NOUN VERB ADJ ADV N+N ADJ+N	0 h

Fig. 7: Your newly created corpora

Merging the two virtual corpora

- In the NOW corpus, click on **Texts/Virtual** in the search form and then click on **Edit Corpora** to get to the list of your virtual corpora shown in Fig. 7.

Fig. 8: Edit corpora

- Now click on the name of the corpus that you want to take the texts from, in this case the “Black Lives Matter” corpus.
- Tick the box before **100** in the light blue bar to select all of the 100 texts in the “Black Lives Matter” corpus. You can also select or deselect individual texts that you wish or do not wish to add to the other corpus.
- Click on the white box that says **—SELECT—** and select the corpus to which you want to add the selected texts, in this case the “Black Lives Matter Floyd” corpus.
- Finally, click on **ADD TO**.

6. You can check if your query has been successful by clicking on **(SEE ALL VIRTUAL CORPORA)** in the blue bar at the top to get back to the overview of your virtual corpora. The number of texts listed in the column **#TEXTS** should have changed now, relative to the number of texts you have added to the corpus.

BLACK LIVES MATTER (RENAME) [DELETE] [ADD TO] [MOVE TO] [--SELECT--] (SEE ALL VIRTUAL CORPORA) [SHARE LIST ?]				
HELP	#TEXTS	DATE	COUNTRY	TITLE
1	<input checked="" type="checkbox"/>	14-12-20	Canada	Black Lives Matter makes its message heard clearly...
2	<input checked="" type="checkbox"/>	16-07-15	Canada	Meet some of the queer members of Black Lives Matt...
3	<input checked="" type="checkbox"/>	20-06-29	Canada	What It's Like to Support Black Lives Matter as a ...
4	<input checked="" type="checkbox"/>	17-10-07	Great Britain	FBI warns of new threat of 'black identity extremi...
5	<input checked="" type="checkbox"/>	18-05-05	Great Britain	'Casspir the friendly ghost': The Apartheid-era we...
6	<input checked="" type="checkbox"/>	18-05-07	Great Britain	More than 100000 years of life lost to police viol...
7	<input checked="" type="checkbox"/>	19-05-07	Great Britain	Sandra Bland: Family requests case be reopened aft...
8	<input checked="" type="checkbox"/>	20-06-08	Great Britain	Around 1,000 people gather for Black Lives Matter ...
9	<input checked="" type="checkbox"/>	20-06-18	Great Britain	Sadiq Khan urges Londoners not to join protests am...
10	<input checked="" type="checkbox"/>	20-07-10	Great Britain	What Black Lives Matter can learn from the 1960s s...
11	<input checked="" type="checkbox"/>	20-10-30	Great Britain	Georgia declares state of emergency and authorises...
12	<input checked="" type="checkbox"/>	20-12-16	Great Britain	Simeon Francis: Investigation launched after black...
13	<input checked="" type="checkbox"/>	20-06-08	India	Global protests over police killing of George Floy...
14	<input checked="" type="checkbox"/>	20-07-22	India	Global wave of protests against racial inequality ...
15	<input checked="" type="checkbox"/>	15-07-19	United States	Democrats lose control of presidential event
16	<input checked="" type="checkbox"/>	15-12-21	United States	Why America Is Moving Left
17	<input checked="" type="checkbox"/>	15-12-28	United States	The rise of Black Lives Matter: Trying to break th...
18	<input checked="" type="checkbox"/>	16-07-13	United States	Opinion Voice of The Southern
19	<input checked="" type="checkbox"/>	16-07-27	United States	Freddie Gray case: Charges dropped against remaini...

Fig. 9: Adding texts to a corpus

3.2 Collecting vocabulary

This section will demonstrate two ways of searching for collocations and patterns using your virtual corpus/corpora. The searches and collocations mentioned here serve as examples and are by no means exhaustive. The instructions should encourage you to look for more potentially relevant language patterns for your lesson, and/or to guide your students through the search for more patterns in their virtual corpus/corpora themselves.

3.2.1 Searching for collocations using the “List”-function

1. In the NOW corpus, click on **List** and type in “white NOUN”.
2. Click on **Options** and group by **LEMMAS**. This way, different forms of the same word will not be counted twice, e.g., *supremacist* and *supremacists*.
3. Click on **Texts/Virtual** and select your virtual corpus.
4. Click on **Find matching strings** to submit your search.

List Chart Collocates Compare KWIC

white NOUN noun.ALL - ▶_pos

Find matching strings Reset

Search by date

Sections Texts/Virtual Sort/Limit Options

HITS 100

KWIC 200

GROUP BY LEMMAS

DISPLAY LEMMAS

SAVE LISTS WORDS

 NONE (SHOW POS)

 BOTH WORDS

 BOTH LEMMAS

Fig. 10: Searching for noun collocates of the adjective "white"

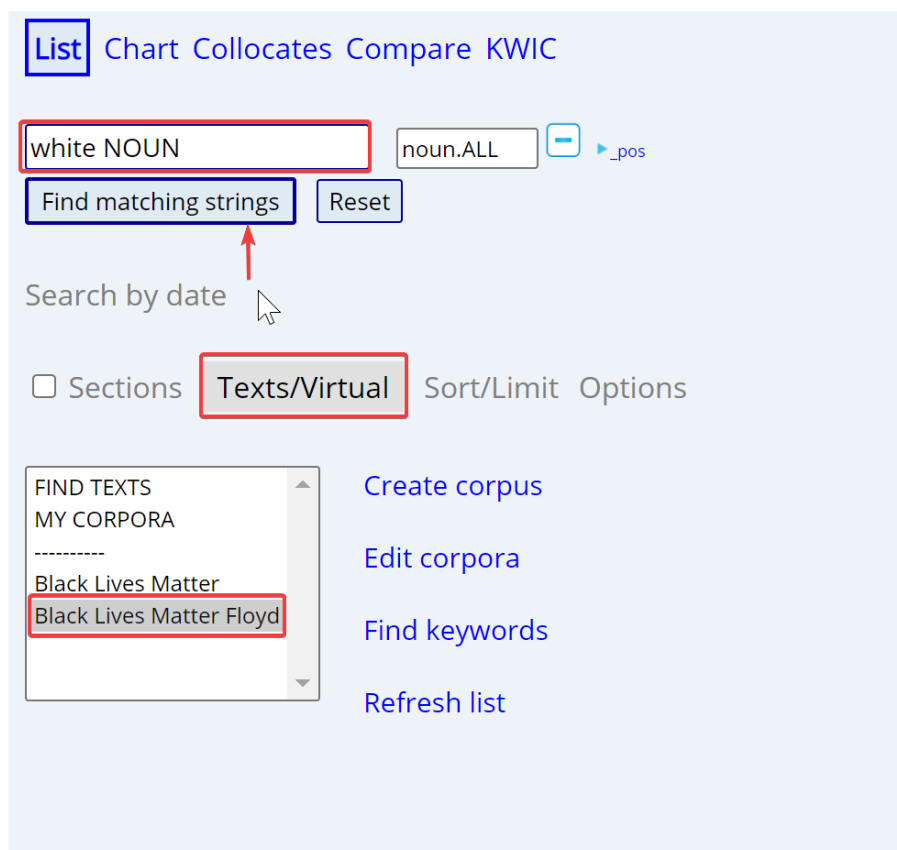


Fig. 11: Selecting the virtual corpus

5. Select collocations that you think are most relevant to the “Black Lives Matter” movement and that you would like to introduce to your students, e.g., *white supremacy*, *white privilege*, *white silence*. Teachers and students are encouraged to highlight the collocations they find most relevant.

ON CLICK: [CONTEXT](#) [TRANSLATE \(??\)](#) [GOOGLE](#) [IMAGE](#) [PRON/VIDEO](#) [BOOK](#) (HELP)

HELP		ALL FORMS (SAMPLE): 100 200 500	FREQ
1	<input type="checkbox"/>	[WHITE] [HOUSE]	395
2	<input type="checkbox"/>	[WHITE] [POLICE]	152
3	<input type="checkbox"/>	[WHITE] [PEOPLE]	78
4	<input type="checkbox"/>	[WHITE] [OFFICER]	61
5	<input type="checkbox"/>	[WHITE] [MAN]	47
6	<input type="checkbox"/>	[WHITE] [SUPREMACIST]	30
7	<input type="checkbox"/>	[WHITE] [SUPREMACY]	27
8	<input type="checkbox"/>	[WHITE] [WOMAN]	24
9	<input type="checkbox"/>	[WHITE] [COP]	18
10	<input type="checkbox"/>	[WHITE] [AMERICAN]	13
11	<input type="checkbox"/>	[WHITE] [POLICEMAN]	10
12	<input type="checkbox"/>	[WHITE] [PERSON]	10
13	<input type="checkbox"/>	[WHITE] [COMMUNITY]	9
14	<input type="checkbox"/>	[WHITE] []	7
15	<input type="checkbox"/>	[WHITE] [FATHER]	7
16	<input type="checkbox"/>	[WHITE] [COUNTERPART]	7
17	<input type="checkbox"/>	[WHITE] [NEIGHBORHOOD]	7
18	<input type="checkbox"/>	[WHITE] [NATIONALIST]	6
19	<input type="checkbox"/>	[WHITE] [PRIVILEGE]	6
20	<input type="checkbox"/>	[WHITE] [SILENCE]	5
21	<input type="checkbox"/>	[WHITE] [FRIEND]	5
22	<input type="checkbox"/>	[WHITE] [COAT]	4

Fig. 12: Search results for “white NOUN”

- Click on the words in the list to see the collocations in context. To get a more balanced overview, generate a random sample by clicking on **100** next to **FIND SAMPLE**. You can use the concordance lines to extract useful example sentences for your students, create a worksheet, or show them a screenshot of the concordance lines to look for authentic contexts and patterns.

FIND SAMPLE: 100 200 500 1000 ←
 PAGE: << < 1 / 264 > >>

CLICK FOR MORE CONTEXT			EXPLORE NEW FEATURES	SAVE	TRANSLATE	ANALYZE
1	20-06-29 --	Smithsonian Magazine	🔍	🔍	🔍	essays reject the notion of " Minnesota Nice, " illustrating how racism and white supremacy pervade people of color's experiences in the
2	17-02-17 US	Chicago Tribune	🔍	🔍	🔍	and SAT scores in applications restricts opportunities for the poor and thus upholds " white supremacy , " # Race relations have been a c
3	20-06-12 US	slate.com	🔍	🔍	🔍	asked at one point. " As long as there's a system of white supremacy there will always be these situations, " DJ Envy replied. At one point
4	15-06-22 US	The Atlantic	🔍	🔍	🔍	. Toward the end of that century, as activists began to effectively challenge white supremacy , its upholders reached for a familiar symbo
5	19-06-23 US	The Hill	🔍	🔍	🔍	provide a space that is inclusive of all and also allow support for open white supremacy , " the statement added. " Support of the Trump
6	20-07-23 IN	morungexpress.com	🔍	🔍	🔍	as a reminder of the enslavement of Black Americans and as a symbol of white supremacy . # Supporters say the flag represents the Sou
7	19-12-26 GB	theguardian.com	🔍	🔍	🔍	. So, if you don't want to deal with racism today and white supremacy today , what you do is you embrace a historical narrative that minir
8	19-02-23 IE	Extra.ie	🔍	🔍	🔍	the toast of Hollywood, claiming a long overdue first Oscar for his 70s-set white supremacy dramedy BlackKlansman, but one person n
9	20-11-05 US	tampabay.com	🔍	🔍	🔍	we have to understand that we, too, are seen as part of white supremacy when it comes to the experience of people of color in this cour
10	20-10-28 US	YAHOO!	🔍	🔍	🔍	. # " The Lee monument was built to celebrate the Confederacy and uphold white supremacy , " Northam said. " This victory moves Virgi
11	20-12-21 US	Salon	🔍	🔍	🔍	The real answer here is we have to acknowledge the reality and threat of white supremacy and racism. The reason why the United State:
12	20-06-12 NG	opinion.premiumtimesng.com	🔍	🔍	🔍	to even surface that argument at all, and as such the systems of white supremacy are highly embedded and rarely challenged. I mean, y
13	20-12-03 US	nytimes.com	🔍	🔍	🔍	with the Department of Homeland Security directed agency analysts to downplay threats from violent white supremacy and Russian elec
14	20-06-16 US	NBC News	🔍	🔍	🔍	" the result of inhumane police brutality that is perpetuated by a culture of white supremacy " and that the officers involved in his death
15	13-11-14 US	TIME	🔍	🔍	🔍	nt think all white people are white supremacist but MORE white people have practiced white supremacy in more places for a longer peri
16	19-08-20 US	sfchronicle.com	🔍	🔍	🔍	new study by the Center for the Study of Hate and Extremism suggests that white supremacy is the most ascendant form of hate around
17	19-03-06 US	Washington Post	🔍	🔍	🔍	whether she believes white supremacy still exists, she shoots back, " Define white supremacy , " and then defines it as " third-wave femir
18	20-06-04 US	Indianapolis Star	🔍	🔍	🔍	above that of anyone who isn't. This, quite simply, is white supremacy ; # How many times have you heard a white person take offense a
19	21-01-13 US	The Boston Globe	🔍	🔍	🔍	of Black lives. The first step in that process is to root out white supremacy , starting with the white supremacist-in-chief. " # Republican Jc
20	20-06-05 US	Indianapolis Star	🔍	🔍	🔍	nearly 200 social media accounts linked to white supremacy groups that planned to encourage members to attend protests over police l
21	21-03-19 US	ABC	🔍	🔍	🔍	also had some advice on confronting racism. # " We need to dismantle white supremacy at every turn, " she said. # " This means looking

Fig. 13: Concordance lines for "white NOUN"

- You can follow the same steps as above to search for adjectives that collocate with *racism*.

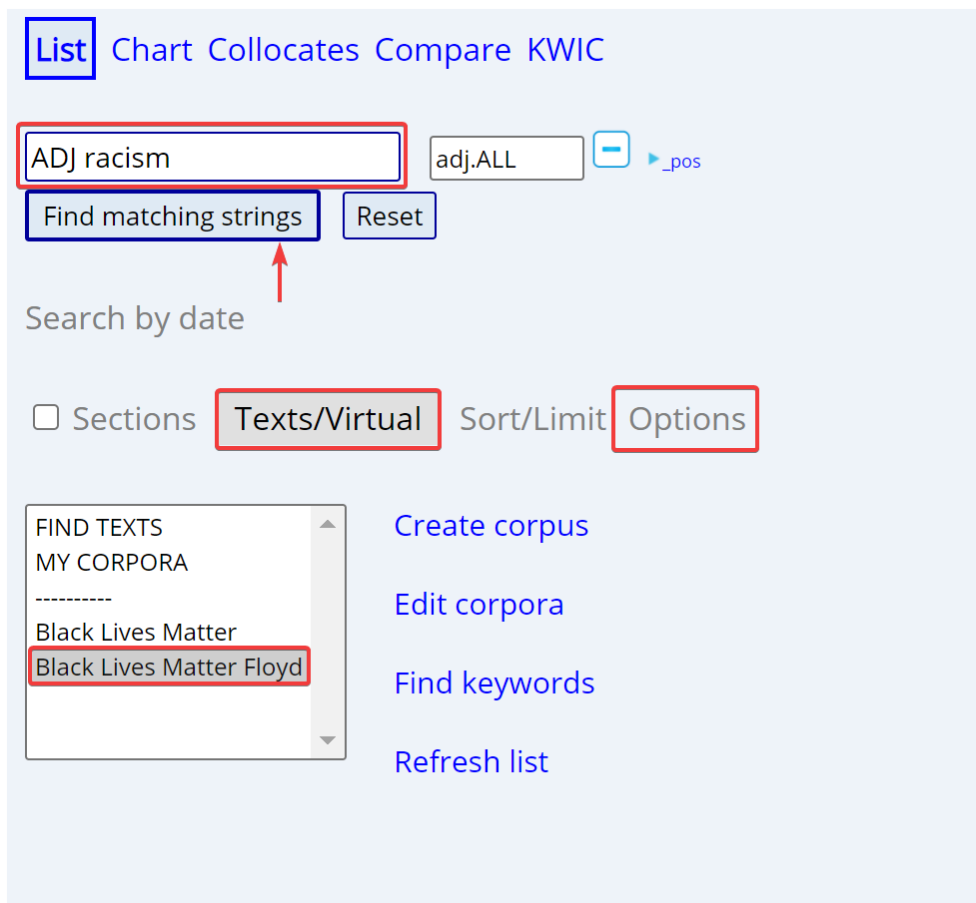


Fig. 14: Searching for adjective collocates of "racism"

ON CLICK: [CONTEXT](#) [TRANSLATE \(??\)](#) [GOOGLE](#) [IMAGE](#) [PRON/VIDEO](#) [BOOK](#) (HELP)

HELP		ALL FORMS (SAMPLE): 100 200 500	FREQ
1	<input type="checkbox"/>	[SYSTEMIC] [RACISM]	135
2	<input type="checkbox"/>	[INSTITUTIONAL] [RACISM]	25
3	<input type="checkbox"/>	[INSTITUTIONALIZED] [RACISM]	12
4	<input type="checkbox"/>	[SYSTEMATIC] [RACISM]	12
5	<input type="checkbox"/>	[STRUCTURAL] [RACISM]	6
6	<input type="checkbox"/>	[INGRAINED] [RACISM]	4
7	<input type="checkbox"/>	[WHITE] [RACISM]	3
8	<input type="checkbox"/>	[LONG-STANDING] [RACISM]	3
9	<input type="checkbox"/>	[INSTITUTIONALISED] [RACISM]	3
10	<input type="checkbox"/>	[COVERT] [RACISM]	2
11	<input type="checkbox"/>	[AMERICAN] [RACISM]	2
12	<input type="checkbox"/>	[SENSELESS] [RACISM]	2
13	<input type="checkbox"/>	[NAKED] [RACISM]	2
14	<input type="checkbox"/>	[STILL] [RACISM]	1
15	<input type="checkbox"/>	[SUSTAINED] [RACISM]	1
16	<input type="checkbox"/>	[SUFFERING] [RACISM]	1
17	<input type="checkbox"/>	[SUBTLE] [RACISM]	1
18	<input type="checkbox"/>	[ONLY] [RACISM]	1
19	<input type="checkbox"/>	[ONGOING] [RACISM]	1
20	<input type="checkbox"/>	[INTERNALIZED] [RACISM]	1
21	<input type="checkbox"/>	[] [RACISM]	1
22	<input type="checkbox"/>	[EVERYDAY] [RACISM]	1
23	<input type="checkbox"/>	[ENDEMIC] [RACISM]	1

Fig. 15: Results for "ADJ racism"

CLICK FOR MORE CONTEXT			EXPLORE NEW FEATURES	SAVE	TRANSLATE	ANALYZE
1	20-06-08 US	denverpost.com	🔍	🔍	🔍	of the coronavirus pandemic, President Donald Trump and decades of police brutality and systemic racism she is not surprised to see protests. #
2	20-06-08 US	denverpost.com	🔍	🔍	🔍	Georgia -- became a rallying point for Americans across the country to call out systemic racism and demand an end to police brutality. # But ther
3	20-06-08 US	aljazeera.com	🔍	🔍	🔍	said. # Evers said Thursday the protests are a watershed opportunity to fix systemic racism . He encouraged people to demonstrate lawfully. # " F
4	20-06-08 US	aljazeera.com	🔍	🔍	🔍	of colour in our nation who rightfully demand an end to 400 years of systemic racism and white supremacy in our country ", Budde said in a state
5	20-06-08 US	nbcnews.com	🔍	🔍	🔍	George Floyd, the Minneapolis man whose death touched off a national debate about systemic racism , were joined by thousands of strangers wf
6	20-06-07 US	washingtonexaminer.com	🔍	🔍	🔍	between law enforcement and the communities they serve as protests against police brutality and systemic racism stretch into day six across the
7	20-06-06 US	thehill.com	🔍	🔍	🔍	to " dominate " the streets to quell protests related to police brutality and systemic racism , took no questions on Friday in the Rose Garden. Whei
8	20-06-06 US	cnbc.com	🔍	🔍	🔍	against black Americans, protesters gathered in cities across the country to stand against systemic racism . # Here's a breakdown of what happen
9	20-06-06 US	nypost.com	🔍	🔍	🔍	to have a moment of silence for every black life that was lost to systemic racism . " 31883647 # Still, she never blocked
10	20-06-05 US	euronews.com	🔍	🔍	🔍	over Floyd's death spread quickly to other major American cities in solidarity against systemic racism in the country and the Black Lives Matter m
11	20-06-05 US	euronews.com	🔍	🔍	🔍	died in an incident that has sparked protests against police violence and allegations of systemic racism . # Floyd died in Minneapolis on May 25 af
12	20-06-05 US	nytimes.com	🔍	🔍	🔍	basis of race, and that his election would not guarantee an end to systemic racism . # " Hate didn't begin with Donald Trump; it's not going
13	20-06-05 US	nytimes.com	🔍	🔍	🔍	that protesters have repeated in recent days, urging attention to the issue of systemic racism and criminal justice reform in nationwide demonst
14	20-06-05 US	columbian.com	🔍	🔍	🔍	the death of George Floyd have revived anger in France over police violence, systemic racism and the complicated case of Adama Traore, a black
15	20-06-05 US	columbian.com	🔍	🔍	🔍	George Floyd's death have revived anger in France over alleged police violence, systemic racism and the complicated case of Adama Traore, a bla
16	20-06-04 US	Patch	🔍	🔍	🔍	brutality, but also reminded the crowd of America's long racial history, systemic racism and multifaceted issues, such as the industrial prison corr
17	20-06-04 US	Time	🔍	🔍	🔍	National Security Adviser Robert O'Brien told CNN he doesn't " think there's systemic racism " in law enforcement in the country. " There's a few l
18	20-06-04 US	abcnews.go.com	🔍	🔍	🔍	vent our rage toward them. Our rage needs to be channeled toward the systemic racism that we're there to protest. " # Davis, who witnessed the
19	20-06-04 US	abcnews.go.com	🔍	🔍	🔍	have been able to keep the focus of the protests on larger issues -- systemic racism and injustice. # Protesters rally against the death in Minneap
20	20-06-04 US	People	🔍	🔍	🔍	way. I am exceptionally proud of you. " # To help combat systemic racism , consider learning from or donating to these organizations: # * Campai
21	20-06-04 US	Bleacher Report	🔍	🔍	🔍	formerNew York YankeespitcherCC Sabathiaand his family joined with others across the city in protesting systemic racism and police brutality fol
22	20-06-04 US	Forbes	🔍	🔍	🔍	Agency via Getty Images # Following protests around the world against police brutality and systemic racism . Ubisoft has cancelled the reveal of t

Fig. 16: Concordance lines for "systemic racism"

Note that some of these words are different types of racism, e.g.:

- *systemic racism*
- *institutional racism*
- *structural racism*
- *covert racism*

You can use these types and also terms like *white supremacy* as topics for further lessons on racism and the "Black Lives Matter" movement and use the corpus to find more language patterns relevant to these topics. To get a better understanding of these forms of racism, you may consult these and other sources:

- <https://www.racialequitytools.org/glossary>
- <http://www.aclrc.com/forms-of-racism>
- <https://theconversation.com/explainer-what-is-systemic-racism-and-institutional-racism-131152>

8. As a final example, these are the results of the query "VERB racism":

HELP		ALL FORMS (SAMPLE): 100 200 500	FREQ
1	<input type="checkbox"/>	[CONDEMN] [RACISM]	13
2	<input type="checkbox"/>	[ADDRESS] [RACISM]	7
3	<input type="checkbox"/>	[DENOUNCE] [RACISM]	6
4	<input type="checkbox"/>	[END] [RACISM]	4
5	<input type="checkbox"/>	[PROTEST] [RACISM]	4
6	<input type="checkbox"/>	[TACKLE] [RACISM]	4
7	<input type="checkbox"/>	[TOLERATE] [RACISM]	3
8	<input type="checkbox"/>	[STOKE] [RACISM]	3
9	<input type="checkbox"/>	[FIGHT] [RACISM]	3
10	<input type="checkbox"/>	[DEFEAT] [RACISM]	3
11	<input type="checkbox"/>	[COMBAT] [RACISM]	2
12	<input type="checkbox"/>	[CHALLENGE] [RACISM]	2
13	<input type="checkbox"/>	[CALL] [RACISM]	2
14	<input type="checkbox"/>	[ENCOUNTER] [RACISM]	2
15	<input type="checkbox"/>	[DISMANTLE] [RACISM]	2
16	<input type="checkbox"/>	[EXPERIENCE] [RACISM]	2
17	<input type="checkbox"/>	[REJECT] [RACISM]	2
18	<input type="checkbox"/>	[SHOW] [RACISM]	2
19	<input type="checkbox"/>	[ERADICATE] [RACISM]	2
20	<input type="checkbox"/>	[UPHOLD] [RACISM]	1
21	<input type="checkbox"/>	[THINK] [RACISM]	1
??	<input type="checkbox"/>	[SUPPORT] [RACISM]	1

Fig. 17: Results for “VERB racism”

Here, students could be asked to classify these collocations according to their meaning. Depending on the students’ proficiency level and previous experience with the topic, they may create their own semantic categories or you can provide them with ready-made categories. For weaker students, an example can be included for each category. In the hands-on alternative to this lesson plan, students can be encouraged to click on these individual collocations to look for example sentences and to take note of the most frequent syntactic patterns in which these collocations occur, e.g., *protest against racism, to dismantle the systems that perpetuate racism, we do not condone racism.*

3.2.2 Searching for patterns in the concordance lines

This method allows you (and your students) to look at both lexical and grammatical patterns in context.

1. In the “List”-function, type “police” into the white box.
2. Click on **Texts/Virtual** and select your virtual corpus.
3. Click on **Find matching strings**.

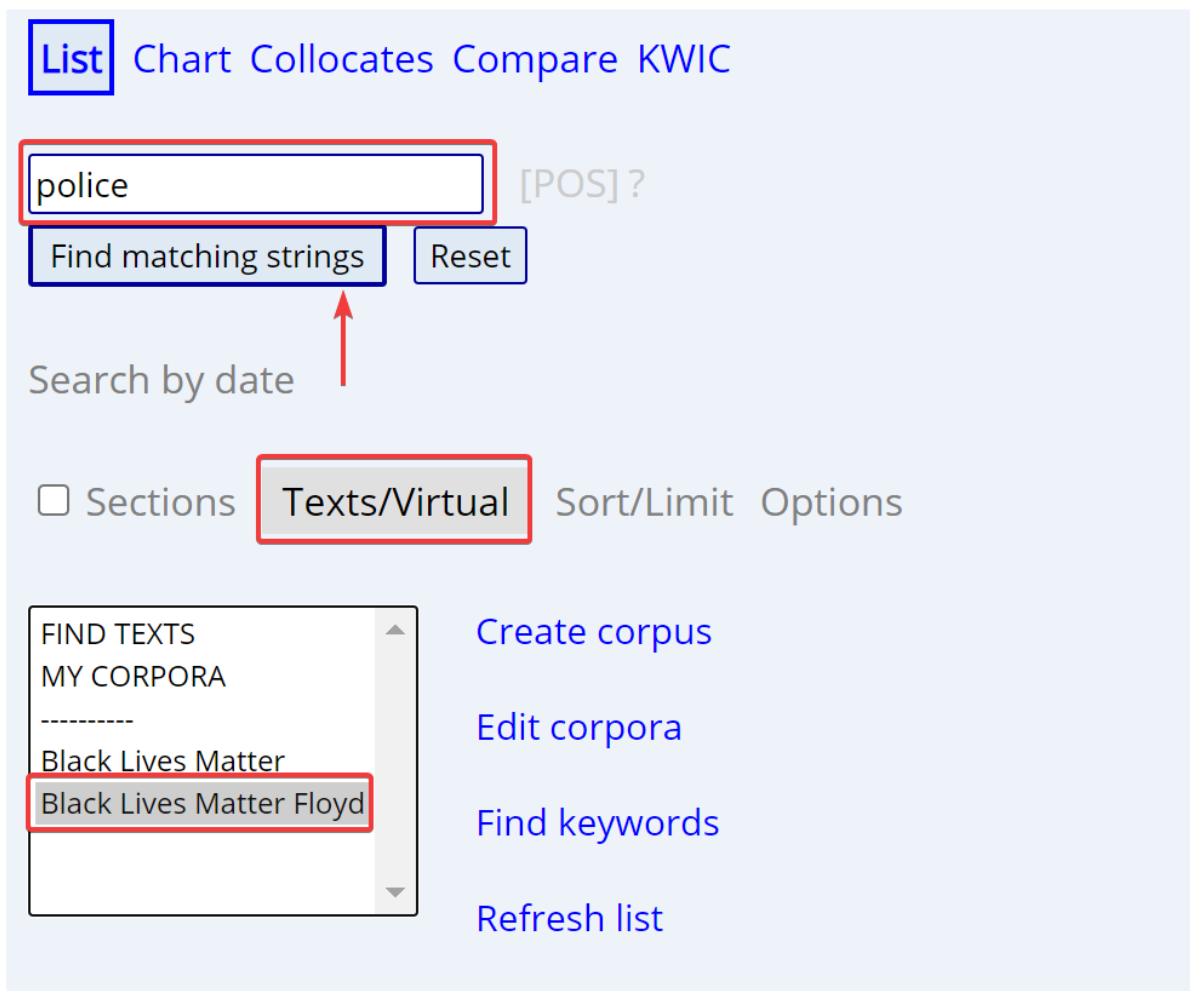


Fig. 18: Search for the word "police"

4. On the next page, click on "POLICE" to get to the concordance lines.
5. Click on **100** next to **Find Sample** to get a random sample.
6. Look through the concordance lines to search for recurrent patterns and/or let your students pick out relevant patterns.

CLICK FOR MORE CONTEXT		EXPLORE NEW FEATURES		SAVE	TRANSLATE	ANALYZE
1	20-05-30 US	reviewjournal.com	🔍🔍🔍	Q	have been erupting all over the country after George Floyd died earlier this week in police custody in Minneapolis. (Tammy Ljungblad/The Kansas City	
2	20-06-01 US	sfchronicle.com	🔍🔍🔍	Q	. The youths entered the store, took "several items" and fled before police officers arrived. Store employees were still determining the loss and no a	
3	20-06-04 US	Chicago Tribune	🔍🔍🔍	Q	to an influx of complaints to Chicago's Civilian Office of Police Accountability. The police disciplinary agency took in 82 complaints from Friday mornir	
4	20-06-08 US	aljazeera.com	🔍🔍🔍	Q	tweeted. "A full 25 minutes before the curfew & w/o provocation, federal police used munitions on peaceful protestors in front of the White House, &	
5	20-06-08 US	houstonchronicle.com	🔍🔍🔍	Q	board, with the ability to subpoena documents and witnesses and to investigate claims against police , including use-of-force complaints; # order the	
6	20-06-04 US	time.com	🔍🔍🔍	Q	the White House. An FBI plane, an Army surveillance plane and a Park Police helicopter circled overhead. # At one point near the White House, prote:	
7	20-06-02 US	usatoday.com	🔍🔍🔍	Q	stay until the curfew was lifted at 6 a.m. Reports on social media showed police corralling a group of demonstrators onto one D.C. street as they use:	
8	20-06-08 US	nytimes.com	🔍🔍🔍	Q	"Mr. Dinkins said. "It could have easily been my body under that police officer. " # -- John Eligon # Oakland, Calif. Qiana Walker	
9	20-06-02 US	YAHOO!	🔍🔍🔍	Q	Instagram, according to the department. The post was taken down, however, police officials said it violated the department's social media policy and	
10	20-05-29 US	chicagotribune.com	🔍🔍🔍	Q	. Four Minneapolis officers involved in the arrest of the black man who died in police custody were fired, hours after a bystander's video showed an c	
11	20-05-28 US	time.com	🔍🔍🔍	Q	following Floyd's death. # After Floyd, who was black, died in police custody, bystander video revealed that a white police officer knelt on his neck at	
12	20-06-08 US	aljazeera.com	🔍🔍🔍	Q	should adopt. " # This, he said, includes "excessive use of police or state force. Given the sensitive history in the US, this deserves special	
13	20-06-03 US	YAHOO!	🔍🔍🔍	Q	donating to these organizations: # *Campaign Zero (joincampaignzero.org) which works to end police brutality in America through research-proven s	
14	20-06-08 US	aljazeera.com	🔍🔍🔍	Q	Donald Trump wants governors to utilise more National Guard troops to respond to protests against police brutality . # Earlier on Monday, Trump tol	
15	20-05-29 US	cbsnews.com	🔍🔍🔍	Q	use of a counterfeit \$20 bill. # Prosecutors said Floyd complied with orders from police officers to leave his vehicle but did not "voluntarily" get in th	
16	20-06-03 US	NPR	🔍🔍🔍	Q	I actually, at one point, got in a public argument with the Minneapolis police chief when he actually took these grand jury considerations, which we h:	
17	20-06-08 US	aljazeera.com	🔍🔍🔍	Q	treatment of Indigenous people. # The protests have split opinion, with some state police and legislators approving the action despite the health risk	
18	20-05-29 US	cbsnews.com	🔍🔍🔍	Q	report from the medical examiner. # JUST IN: Derek Chauvin, the former police officer who knelt on George Floyd's neck, has been charged with th	
19	20-06-04 US	wbur.org	🔍🔍🔍	Q	was a short but very successful experiment in so-called community policing, where we had police in sufficient numbers out on the street, building rel	
20	20-06-04 US	YAHOO!	🔍🔍🔍	Q	more than eight minutes in Minneapolis, Minnesota. On Wednesday, three more former police officers were charged in connection with his death -- f	
21	20-06-08 US	aljazeera.com	🔍🔍🔍	Q	in Washington, DC Jim Bourg/Reuters # Protesters kneel in front of New York City Police during a march to honour George Floyd in New York City Joh	
22	20-06-02 US	Washington Times	🔍🔍🔍	Q	. 85082020 Question of the Day Should police do more to curtail rioting? Story TOPics # Police and National Guard block	
23	20-06-01 US	sfchronicle.com	🔍🔍🔍	Q	with vandalized windows and property in downtown Walnut Creek, according to the Walnut Creek police department. The city's curfew took effect at	

Fig. 19: Concordance lines for "police"

- In this example, students may be encouraged to pick out compound nouns like *police brutality*, *police officer*, collocations like *to protest against police brutality* or *to investigate claims against police*. You can also point out that *police* appears sometimes with and sometimes without an article, as you can see in the two examples *killed at the hands of police* and *the police in his city have arrested* (Fig. 20).

56	20-06-03 US	mercurynews.com	🔍🔍🔍	Q	we have to call for -- we called out our National Guard and our street police , but the rhetoric that's coming out of the White House is making it
57	20-06-03 US	Us Weekly	🔍🔍🔍	Q	not have that anymore. " # Floyd was killed on May 25 when a police officer pressed his knee into the Minneapolis resident's neck for nearly nine
58	20-05-29 US	baltimoresun.com	🔍🔍🔍	Q	She called for a civilian review board that would be empowered to take action on police complaints. She added that such efforts must be paired
59	20-06-04 US	Fox News	🔍🔍🔍	Q	discuss policing and the civil unrest that has followed the killing of George Floyd by police in Minneapolis. (My Brother's Keeper Alliance and The
60	20-05-28 US	twincities.com	🔍🔍🔍	Q	was shot multiple times by a man with a criminal history that includes interfering with police and harassing public officials, authorities said Thur:
61	20-06-03 US	euronews.com	🔍🔍🔍	Q	of US law enforcement. She says much more needs to be done to increase police accountability . 'Sowing discord' # Foes and critics of the Unite
62	20-06-01 US	seattletimes.com	🔍🔍🔍	Q	of George Floyd, a Black man who died after being pinned under a white police officer's knee pressed to his throat. # Floyd's killing " is bad
63	20-06-03 US	Milwaukee Journal Sentinel	🔍🔍🔍	Q	# On Memorial Day, an unarmed African American was killed at the hands of police . Watching the video of the death of George Floyd is heart-wr
64	20-05-27 US	nytimes.com	🔍🔍🔍	Q	the spokesman, John Elder. # During protests on Tuesday, some people vandalized police vehicles with graffiti and targeted the precinct house v
65	20-05-26 US	Billboard	🔍🔍🔍	Q	George Floyd on video as witnesses begged the police officer to take him into the police car and get off his neck," said Ben Crump, a civil rights
66	20-06-01 US	cbsnews.com	🔍🔍🔍	Q	our calculations. # As you know, there are calls for the other three police officers involved in George Floyd's arrest and subsequent death to also
67	20-06-04 US	Time	🔍🔍🔍	Q	Chauvin's case -- from his long list of previous conduct complaints to the 44-year-old police officer's brutal calm as he pressed his knee to Georg
68	20-06-03 US	nytimes.com	🔍🔍🔍	Q	. " # Image # Credit... Erin Schaff/The New York Times # U.S. Park Police officers pushed protesters back near the White House. # Image # Credit
69	20-06-05 US	arkansasonline.com	🔍🔍🔍	Q	report. # Hall told Good Morning America that the situation escalated quickly and that police grabbed Floyd, put him in a squad car, dragged him
70	20-06-04 US	ABC	🔍🔍🔍	Q	protesters are demanding justice for Floyd and decrying the overall treatment of Black Americans by police . # The Minneapolis Police Departme
71	20-06-07 US	cbsnews.com	🔍🔍🔍	Q	where a tense encounter with the police ensued. # Protesters kneel in front of police during a demonstration on Broadway near New York City's
72	20-06-03 US	nydailynews.com	🔍🔍🔍	Q	Four cops could've been killed last night when Molotov cocktails were thrown into a police car they were in. Other officers were seriously injured
73	20-06-01 US	sfchronicle.com	🔍🔍🔍	Q	heads from more than a dozen jurisdictions across the county. # 7:07 p.m. Police report arrests, injuries in downtown Walnut Creek as curfew gr
74	20-06-02 US	My San Antonio	🔍🔍🔍	Q	family of Adama Traor?, a 24-year-old black construction worker who died in French police custody in 2016 - the protest attempted to highlight p
75	20-06-04 US	abcnews.go.com	🔍🔍🔍	Q	woman holds a megaphone during a protest against the death of George Floyd in Minneapolis police custody at Brandenburg Gate in Berlin, Gei
76	20-06-03 US	washingtonpost.com	🔍🔍🔍	Q	. For Trump, that figure was 40 percent. # A retired St. Louis police captain and municipal chief whose death was caught on Facebook Live was tr
77	20-06-03 US	nymag.com	🔍🔍🔍	Q	Mayor Eric Garcetti has called for the National Guard; since protests began, the police in his city have arrested motorists at random and fired rut
78	20-05-29 US	chicagotribune.com	🔍🔍🔍	Q	department veteran who has been involved in police shootings over the years. # The police house cats don't have complaints in their files. They
79	20-06-03 US	washingtonpost.com	🔍🔍🔍	Q	unrest stood in contrast to the daylight hours when protesters marched, stopping at Milwaukee police headquarters and arriving around 5 p.m.
80	20-06-01 US	npr.org	🔍🔍🔍	Q	a cellphone video captured some of Eric Garner's final words as New York City police officers sat on his head and pinned him to the ground on a

Fig. 20: Concordance lines for "police"

3.3 Creating Materials

You can create **crossword puzzles** based on your findings from the corpus. An easy way to do this is by using the website linked below. Its use is very easy and self-explanatory. <https://printablecreative.com/crossword-generator>

Word clouds are a very good method to visualize the results drawn from corpora. You can create them very easily using **PowerPoint**. Click on the link below and follow the instructions:

<https://www.bettercloud.com/monitor/the-academy/create-word-cloud-PowerPoint-presentation/>

Additional Advice: It is better to convert your whole text to lowercase letters because sometimes the program differentiates between words that begin with a capital and ones beginning with a lowercase letter. So that it may create two entries for the same word in your word cloud.

4 Lesson plan

Note that this lesson deals with very sensitive issues (e.g., murder, violence, racism and other forms of discrimination) and may therefore require a trigger warning.

Warm-up: (10 min)

1. Write the name "George Floyd" on the blackboard and/or show a picture.
2. Ask the students if they have heard of this person.
3. Ask about the way he died and why they think he died.
4. Write out the name of the "Black Lives Matter" movement.
5. Ask the students if/what they know about it.

Pre-task: (20 min)

1. In pairs or groups of three, tell the students to open their web browser and connect to <https://www.dictionary.com/>.
2. Ask students to type "Black Lives Matter" into the search box (you could also provide the definition on a worksheet).

A political and social movement originating among African Americans, emphasizing basic human rights and racial equality for Black people and campaigning against various forms of racism. BLM, B.L.M.

3. Tell students to read the definition and discuss it in their groups. Ensure that the students understand the main meaning of the definition. You can also ask them to explain it in their mother tongue.
4. Introduce the lesson's aim, which is to learn language patterns related to the "Black Lives Matter" movement and racism in order to be able to speak and write about the topic.

After this point, the progression of the lesson depends on whether you choose to let the students work with the corpus themselves or not.

Option 1 (teacher-corpus interaction only):

Task 2 (30 min.)

1. Tell the students that you will show them lists of words, patterns and sentences collected from different texts on the internet on the topic of "Black Lives Matter".
2. Share the worksheet(s) you created on the basis of your corpus or the screenshots that you made from your collected data.
3. Let your students work with and discuss the materials in pairs or groups.

Task 3 (30 min.)

In this lesson version, you will probably have more time to discuss the results and take a closer look at the various aspects that occur in the material, e.g., the types of racism, the meaning of expressions like *white supremacy* and *white privilege*, etc. You might want to bring in more information on (one of) these aspects and work on them in more detail, using the language patterns introduced in the previous task.

Option 2 (with hands-on student-corpus interaction):

Task 2: (45 min)

1. Guide your students through the creation of a virtual corpus, following the steps described in section 3.1.
2. Show your students how to search for collocations and language patterns in their virtual corpus.
3. Encourage students to search for a number of collocations themselves and discuss the results first in pairs and then in plenum. Collect language patterns that you and your class deem most relevant to talk about the "Black Lives Matter" movement and racism more generally in English.

Task 3: (15 min)

For the last part of your lesson, you can select some exercises of the worksheet below to focus on individual aspects in more detail. Alternatively, you can also use the worksheet activities as homework.

5 Worksheet

The language patterns and sentences used in this worksheet are taken from the virtual corpus. The activities aim to ensure the understanding of the different terms and their uses. It is up to you to decide if your students need an additional glossary to understand the topics on the worksheet (see links above for suitable resources).

A) Complete the crossword below with the help of the two screenshots and the definitions below. You can find the answers to the crossword either in the list of *white* + NOUN or in the list of ADJECTIVE + *racism* collocations.

HELP		ALL FORMS (SAMPLE) : 100 200 500
1	<input type="checkbox"/>	[WHITE] [HOUSE]
2	<input type="checkbox"/>	[WHITE] [POLICE]
3	<input type="checkbox"/>	[WHITE] [PEOPLE]
4	<input type="checkbox"/>	[WHITE] [OFFICER]
5	<input type="checkbox"/>	[WHITE] [MAN]
6	<input type="checkbox"/>	[WHITE] [SUPREMACIST]
7	<input type="checkbox"/>	[WHITE] [SUPREMACY]
8	<input type="checkbox"/>	[WHITE] [WOMAN]
9	<input type="checkbox"/>	[WHITE] [COP]
10	<input type="checkbox"/>	[WHITE] [AMERICAN]
11	<input type="checkbox"/>	[WHITE] [POLICEMAN]
12	<input type="checkbox"/>	[WHITE] [PERSON]
13	<input type="checkbox"/>	[WHITE] [COMMUNITY]
14	<input type="checkbox"/>	[WHITE] []
15	<input type="checkbox"/>	[WHITE] [FATHER]
16	<input type="checkbox"/>	[WHITE] [COUNTERPART]
17	<input type="checkbox"/>	[WHITE] [NEIGHBORHOOD]
18	<input type="checkbox"/>	[WHITE] [NATIONALIST]
19	<input type="checkbox"/>	[WHITE] [PRIVILEGE]
20	<input type="checkbox"/>	[WHITE] [SILENCE]
21	<input type="checkbox"/>	[WHITE] [FRIEND]
22	<input type="checkbox"/>	[WHITE] [COAT]

Fig. 21: Most common noun collocates of “white”

HELP		ALL FORMS (SAMPLE) : 100 200 500
1	<input type="checkbox"/>	[SYSTEMIC] [RACISM]
2	<input type="checkbox"/>	[INSTITUTIONAL] [RACISM]
3	<input type="checkbox"/>	[INSTITUTIONALIZED] [RACISM]
4	<input type="checkbox"/>	[SYSTEMATIC] [RACISM]
5	<input type="checkbox"/>	[STRUCTURAL] [RACISM]
6	<input type="checkbox"/>	[INGRAINED] [RACISM]
7	<input type="checkbox"/>	[WHITE] [RACISM]
8	<input type="checkbox"/>	[LONG-STANDING] [RACISM]
9	<input type="checkbox"/>	[INSTITUTIONALISED] [RACISM]
10	<input type="checkbox"/>	[COVERT] [RACISM]
11	<input type="checkbox"/>	[AMERICAN] [RACISM]
12	<input type="checkbox"/>	[SENSELESS] [RACISM]
13	<input type="checkbox"/>	[NAKED] [RACISM]
14	<input type="checkbox"/>	[STILL] [RACISM]
15	<input type="checkbox"/>	[SUSTAINED] [RACISM]
16	<input type="checkbox"/>	[SUFFERING] [RACISM]
17	<input type="checkbox"/>	[SUBTLE] [RACISM]
18	<input type="checkbox"/>	[ONLY] [RACISM]
19	<input type="checkbox"/>	[ONGOING] [RACISM]
20	<input type="checkbox"/>	[INTERNALIZED] [RACISM]
21	<input type="checkbox"/>	[] [RACISM]
22	<input type="checkbox"/>	[EVERYDAY] [RACISM]
23	<input type="checkbox"/>	[ENDEMIC] [RACISM]

Fig. 22: Most frequent adjective collocates of "racism"

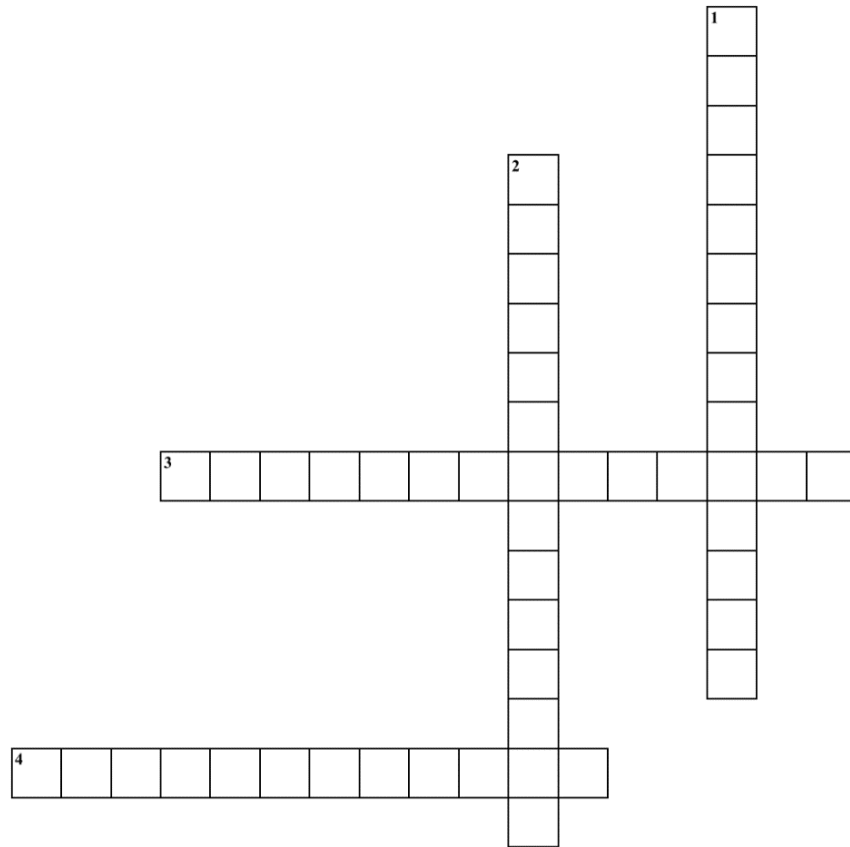


Fig. 23: Crossword puzzle created with <https://printablecreative.com/crossword-generator>

Crossword clues.

1. The idea that white people and the ideas, thoughts, beliefs, and actions of white people are superior to People of Color and their ideas, thoughts, beliefs, and actions. While it is often associated with extremist groups like the Ku Klux Klan and neo-Nazis, it is in fact ever present in our institutional and cultural assumptions.
2. A system in which public policies, institutional practices, cultural representations, and other norms work in different ways to perpetuate racial group inequity. Different aspects of our history and culture have allowed privileges associated with “whiteness” and disadvantages associated with “color” to continue and adapt over time. It is not something that a few people or institutions choose to practice. Instead, it is a feature of the social, economic, and political systems in which we all exist.
3. It refers to the unquestioned and unjustified advantages, benefits and choices that some people enjoy only because they are white. Generally, white people who experience this are unaware of it.
4. A form of racial discrimination that is subtler and often difficult to observe. It is often hidden in the fabric of society, where racist actions are either passive or ambiguous. These subtle actions

are often justified with explanations that society is more willing to believe.

B) Complete the sentences below with the help of the list of verbs that collocate with *racism* below:

HELP		ALL FORMS (SAMPLE) : 100 200 500
1	<input type="checkbox"/>	[CONDEMN] [RACISM]
2	<input type="checkbox"/>	[ADDRESS] [RACISM]
3	<input type="checkbox"/>	[DENOUNCE] [RACISM]
4	<input type="checkbox"/>	[END] [RACISM]
5	<input type="checkbox"/>	[PROTEST] [RACISM]
6	<input type="checkbox"/>	[TACKLE] [RACISM]
7	<input type="checkbox"/>	[TOLERATE] [RACISM]
8	<input type="checkbox"/>	[STOKE] [RACISM]
9	<input type="checkbox"/>	[FIGHT] [RACISM]
10	<input type="checkbox"/>	[DEFEAT] [RACISM]
11	<input type="checkbox"/>	[COMBAT] [RACISM]
12	<input type="checkbox"/>	[CHALLENGE] [RACISM]
13	<input type="checkbox"/>	[CALL] [RACISM]
14	<input type="checkbox"/>	[ENCOUNTER] [RACISM]
15	<input type="checkbox"/>	[DISMANTLE] [RACISM]
16	<input type="checkbox"/>	[EXPERIENCE] [RACISM]
17	<input type="checkbox"/>	[REJECT] [RACISM]
18	<input type="checkbox"/>	[SHOW] [RACISM]
19	<input type="checkbox"/>	[ERADICATE] [RACISM]
20	<input type="checkbox"/>	[UPHOLD] [RACISM]
21	<input type="checkbox"/>	[THINK] [RACISM]

Fig. 24: Most frequent verb collocates of "racism"

Fill in the gaps.

- a) Demonstrations across the US to _____ racism and police abuses remained large but mostly without the violence of previous nights.
- b) This generation's challenge is to _____ racism, which in some ways is even harder. Right? But it can be done.
- c) Thousands of people in Philadelphia continue to _____ racism and police brutality following the death of George Floyd.
- d) "In my life, I have _____ racism," Frahm, 17, said. "It's a terrible feeling."
- e) Europe identified with the cause of US protests and urged their own governments to _____ racism and police violence.

c) Below, you see a word cloud, which shows words commonly related to the word *racism*. The larger the word, the more often it co-occurs with *racism*.



Fig. 25: Word cloud of words associated with "racism" in the BLM corpus (created with PowerPoint)

Words associated with racism

- a) Look at the word cloud. Write down three of the largest words.

1. _____
2. _____
3. _____

b) Write down the meaning of these words in German. You may consult a dictionary for help.

1. _____
2. _____
3. _____

c) Choose three verbs from the word cloud and write three sentences related to our topic.

1.

2.

3.

D) Now look at the word clouds below.

Which topic belongs to which word cloud?

1. George Floyd's death
2. Types of racism
3. Verbs often associated with *racism*
4. The "Black Lives Matter" movement



describes Floyd's death and its context in a very detailed way. The site also provides an audio file of this article, which you can use as a listening activity.

If you wish to add some videos to your lesson to highlight individual aspects and concepts like *white silence* or use them for a follow-up lesson on this topic, here are some links to suggested videos:

White silence on social media:

<https://www.cbsnews.com/news/white-silence-on-social-media-why-not-saying-anything-is-actually-saying-a-lot/>

How *white silence* fuels racism:

<https://www.colorlines.com/articles/icymi-animated-short-your-black-friend-breaks-down-how-white-silence-fuels-racism>

"Black Lives Matter" vocabulary:

<https://www.youtube.com/watch?v=9FOpvZdd35U>

7 Conclusion

Race and racism are relevant topics that ought to be discussed in school. Teachers should be aware that students have their own beliefs and stereotypes, as we all have. At the same time, it is important to show that we live in a biased world and are all susceptible to false beliefs. Maybe some parents or even staff members will find that the topic should not be brought into the classroom. Even if a class mostly consists of white students, this topic is very much justified as part of a public discussion and concerns every citizen in the global, multicultural community that is our world.

The main advantage of planning a lesson using a corpus is that it allows teachers and students to examine language in authentic contexts. Although collecting data can be time consuming, the material collected can serve different purposes in your students' learning process. By following the instructions in the step-by-step guide in section 3, you can create a virtual corpus on a topic of your choice with only a few clicks. Since the NOW corpus is updated on a daily basis, it allows teachers and students to explore issues of topical interest and current affairs from across the world that cannot be found in textbooks.

8 Resources and references

Friginal, Eric. 2018. *Corpus linguistics for English teachers: New tools, online resources, and classroom activities*. New York, NY: Routledge.

Niedersächsisches Kultusministerium. 2017. Kerncurriculum für das Gymnasium, gymnasiale Oberstufe: Englisch. https://cuvo.nibis.de/cuvo.php?k0_0=Schulbereich&v0_0=Sek+II&k0_1=Fach&v0_1=Englisch&docid=1132&p=detail_view (28 June, 2021).

NOW Corpus (News on the Web). <https://www.english-corpora.org/now/>.

Sommier, Mélodine & Roiha, Anssi. 2018. Dealing with Culture in Schools: A small-step approach towards anti-racism in Finland. In Aminkeng A. Alemanji (ed.), *Antiracism education in and out of schools*. Helsinki: Palvagre Macmillan.

Websites:

<https://www.dictionary.com/>.

<https://www.cbsnews.com/news/white-silence-on-social-media-why-not-saying-anything-is-actually-saying-a-lot/>.

<https://www.adl.org/education/resources/tools-and-strategies/how-should-i-talk-about-race-in-my-mostly-white-classroom>.

<https://www.colorlines.com/articles/icymi-animated-short-your-black-friend-breaks-down-how-white-silence-fuels-racism>.

<https://www.newyorker.com/news/daily-comment/the-death-of-george-floyd-in-context>.

<https://printablecreative.com/crossword-generator>.

<https://www.bettercloud.com/monitor/the-academy/create-word-cloud-PowerPoint-presentation/>.

<https://www.racialequitytools.org/glossary#structural-racism>.

<https://www.shorelineschools.org/site/default.aspx?PageType>.

11. Learning words and collocations independently: Rising temperatures and climate science

LISA HÖRNSCHEMEYER AND LENA JANZEN

1 Introduction and rationale

Not only do English as a Foreign Language (EFL) learners need to develop communicative competence, they should also learn how to expand their knowledge independently. The English curriculum for higher level secondary schools in Lower-Saxony, Germany, states that students should add topic-specific vocabulary to their basic knowledge (Niedersächsisches Kultusministerium 2017: 18). Further, students should develop language learning awareness, hence the ability to learn and access new vocabulary using different methods and materials (Niedersächsisches Kultusministerium 2017: 23). Even younger students at lower secondary schools are expected to learn to work with dictionaries and similar resources and then organise new words to expand existing language structures and patterns (Niedersächsisches Kultusministerium 2015: 27). There is a huge variety of resources to look up new words and expressions, but especially non-native speakers may end up using vocabulary in inappropriate contexts. Such problems can be avoided by using authentic language data, which can be found in representative language corpora. Practicing using different resources is often neglected at school or considered “boring” and students frequently only learn vocabulary from their textbook or rely on bilingual dictionaries. By providing learners with rich input and allowing them to discover specific terms and phrases about topics that directly affect and influence them, a teacher can improve individual learning outcomes.

Recently, young people have become more interested and politically active, particularly with regard to the issue of climate change. After Swedish activist Greta Thunberg started protesting in August 2018, millions of students around the world joined her to raise awareness of the climate crisis and set off an avalanche of grassroot organisations demanding political action (Gill 2019). Students should be supported and given the abilities to communicate their opinions and ideas on this topic and the ability to consult corpora can support them in this endeavour by providing them with real-life, relevant language input. The corpus-based lesson presented in this chapter aims at connecting methodological skills and vocabulary training with an affectively and cognitively engaging topic. Teachers can easily adjust the lesson plan to any topic of choice that they consider to be suitable and appealing to their students.

The proposed lesson targets learners at higher secondary school aged 14-15 or older (9th grade). As the students will work online with a corpus, it is assumed that they can use a computer and that they have access to the COCA (Corpus of Contemporary American English) via english-corpora.org. It is best for you to acquire an academic premium license to be able to work with texts longer than 400 words. However, this is not obligatory for this lesson, either. Each student will need to create a free account before the lesson. A handout will guide them through the tasks including step-by-step instructions for the website (cf. Appendix). However, students still have the freedom to search for vocabulary and collocations they consider relevant. Finally, each student will create their own collection of new words and collocations for future use.

Outline and objectives

Teacher preparation time: 30 minutes

Lesson time: 90 minute

Learning objectives: By the end of this lesson, students will be able to:

- access new vocabulary on their own.
- use the COCA to find out about new words and phrases related to a specific topic.
- use phrases in appropriate co- and contexts.
- use methods (e.g., mind-mapping) to structure and store linguistic resources.
- use known and new words and phrases on the topic of climate change.

2 Corpus, tools and methods

In this lesson, students will work with the Corpus of Contemporary American English (COCA), which currently contains texts from 1990 to 2019 and is regularly updated. The COCA can be accessed via english-corpora.org.

3 Step-by-step guide

In this chapter, step-by-step instructions are provided for both the teacher and the students. The teacher will have to provide the students with an appropriate source, e.g., an online article, and adjust the worksheet and handout for the students to match the chosen topic. With the help of a handout with step-by-step instructions, students will work with the corpus themselves.

3.1 Teacher preparation

First of all, as a teacher, skim the lesson plan (Table 7) to have an overview of the planned lesson. Be aware that this lesson should be preceded by an introductory lesson on climate change. For this chapter's lesson, choose an appropriate text or article for the learners that serves as a starting point. This text then needs to be shared with the students either digitally or as a print version. To prepare the text, follow the instructions below:

3.1.1 Text preparation

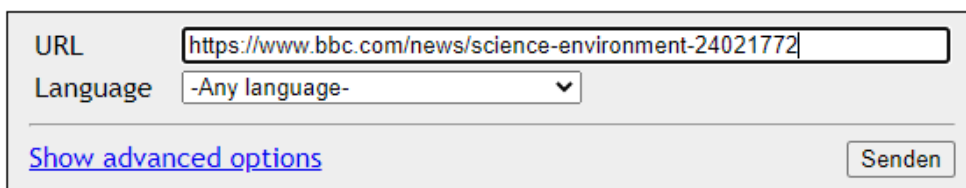
1. Choose a text on the topic you want your students to focus on and derive new vocabulary from. We recommend using Ecosia (<https://www.ecosia.org/>) for this, the search engine that plants trees. We chose

the article “What is climate change? A really simple guide” (<https://www.bbc.com/news/science-environment-24021772> (BBC News 2020).

2. Instead of – or in addition to – providing your students with a digital/print version of the article, you could also make use of the text analysis function of the COCA, which lets you find out more about the words and phrases in your text (e.g., frequencies, collocational behaviour, etc.). If you wish to do so, you need to upload your text onto the website. First, remove any possibly distracting content. To do so, go to “jusText demo” (Pomikalek 2021, URL: <https://nlp.fi.muni.cz/projects/justext/>), copy and paste the link that belongs to the article (Fig. 1).

jusText demo

jusText is a heuristic based boilerplate removal tool. See the project page at [Corpus tools](#).



URL

Language

[Show advanced options](#)

(c) Jan Pomikalek <xpomikal@fi.muni.cz>, [NLPlab](#), FI MU

Fig. 1: Paste URL of source text

3. Press ENTER on your keyboard and select “Hide boilerplate” on your screen (Fig. 2).



jusText demo

jusText is a heuristic based boilerplate removal tool. See the project page at [Corpus tools](#).

URL

Language

[Show advanced options](#)

[Hide boilerplate](#)

BBC News

Fig. 2: Boilerplate removal with jusText

4. Your text will now appear without any photos, images or other annotations (Fig. 3) and you can provide it to your students by copying it into a new document.

jusText demo

jusText is a heuristic based boilerplate removal tool. See the project page at [Corpus tools](#).

URL

Language

[Show advanced options](#)

[Show boilerplate](#)

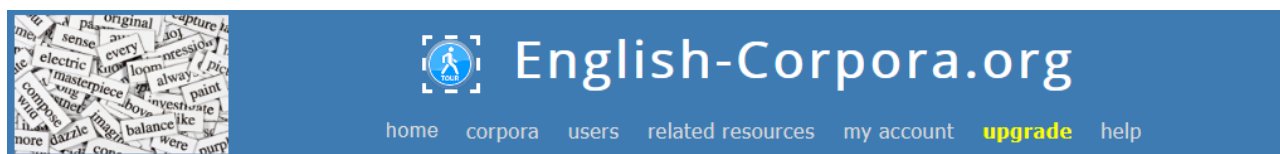
What is climate change? A really simple guide
 While Covid-19 has shaken much of human society, the threat posed by global warming has not gone away.
 Human activities have increased carbon dioxide emissions, driving up temperatures. Extreme weather and melting polar ice are among the possible effects.

What is climate change?
 The Earth's average temperature is about 15C but has been much higher and lower in

Fig. 3: Cleaned text

5. Share the text digitally with your students (e.g., via e-mail, Google Docs or the open source tool Cryptpad). You can also hand out a printed version.

6. In order to prepare the beginning of the lesson, go to www.english-corpora.org/ and select the COCA from the list of available corpora (Fig. 4).



The most widely used online corpora. [Guided tour](#), [overview](#), [search types](#), [variation](#), [virtual corpora](#), [corpus-based resources](#).

The links below are for the online interface. But you can also download the corpora for use on your own computer.

Corpus (online access)	Download	# words	Dialect	Time period	Genre(s)
iWeb: The Intelligent Web-based Corpus		14 billion	6 countries	2017	Web
News on the Web (NOW)		11.9 billion+	20 countries	2010-yesterday	Web: News
Global Web-Based English (GloWbE)		1.9 billion	20 countries	2012-13	Web (incl blogs)
Wikipedia Corpus		1.9 billion	(Various)	2014	Wikipedia
Corpus of Contemporary American English (COCA)		1.0 billion	American	1990-2019	Balanced
Coronavirus Corpus		844 million+	20 countries	Jan 2020-yesterday	Web: News
Corpus of Historical American English (COHA)		400 million	American	1810-2000	Balanced

Fig. 4: List of corpora

If you want to work with a text longer than 400 words, make sure your account icon is not red but green (Fig. 5). Red would mean you are not logged in with your institution's academic or premium account.

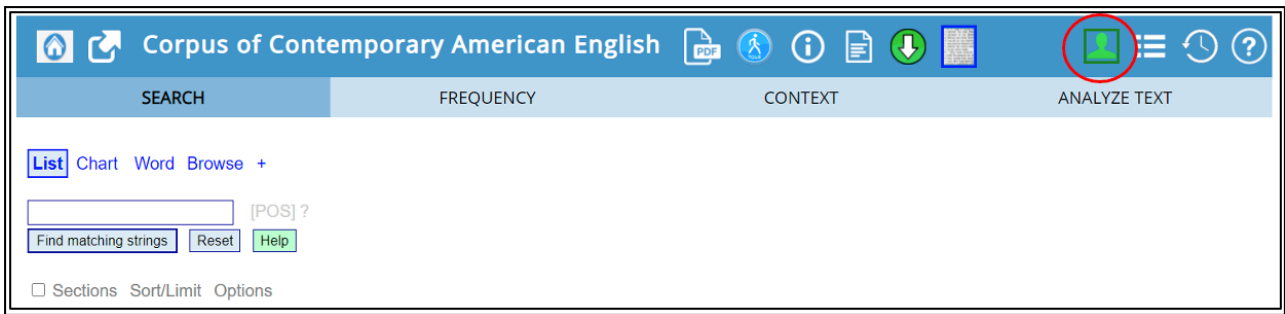


Fig. 5: Log-in icon on english-corpora.org

7. Choose the function “Analyze text” (Fig. 6).

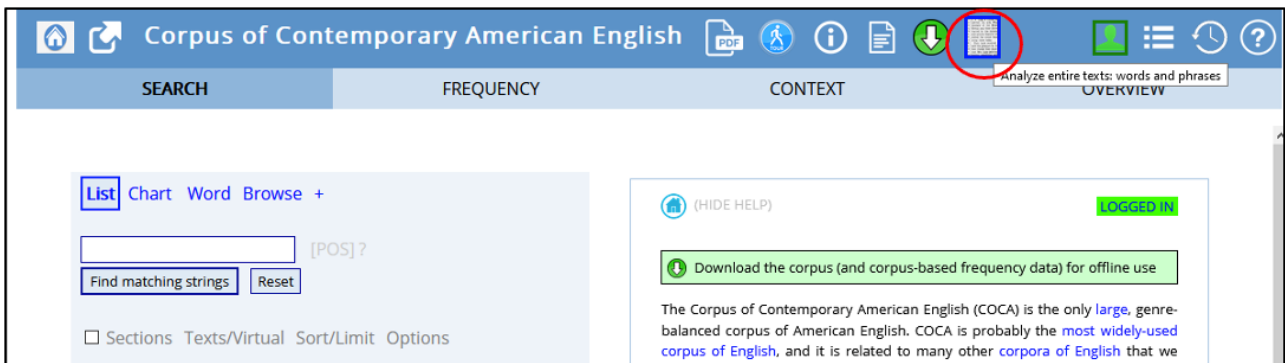


Fig. 6: Icon for the function “Analyze text”

The following interface will appear (Fig. 7).

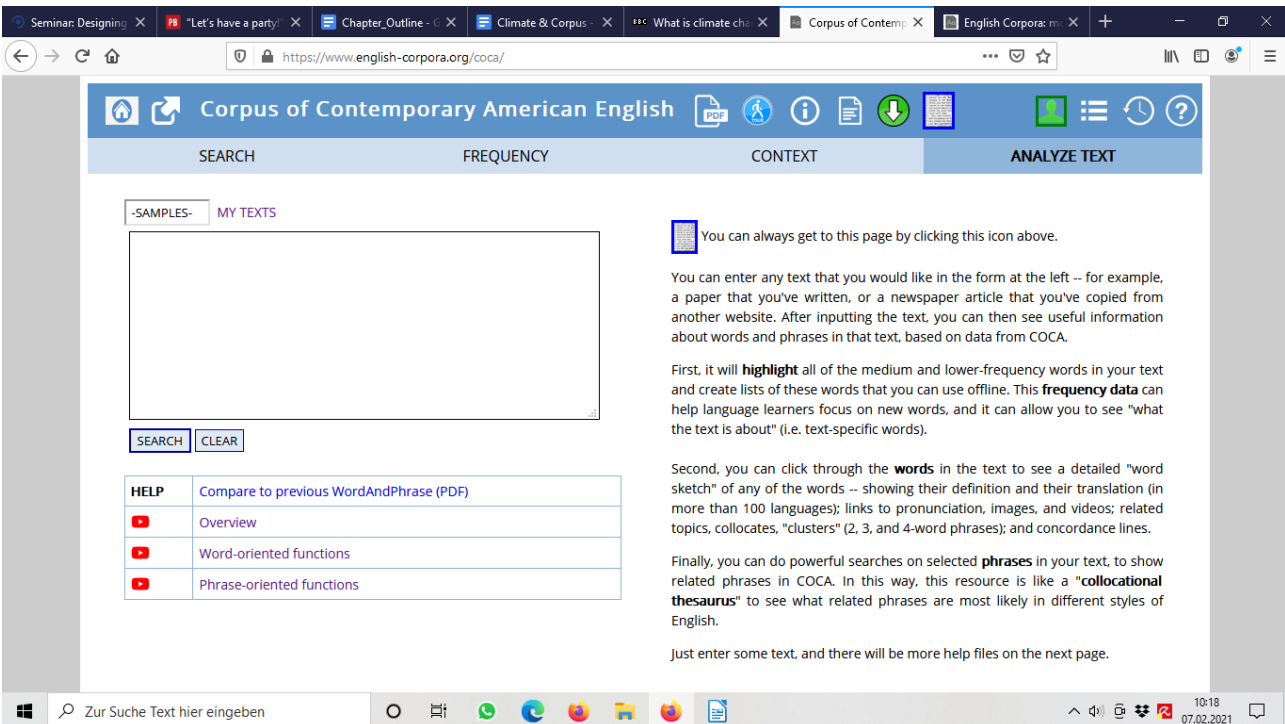


Fig. 7: “Analyze text” – interface

8. Copy the text you chose for this exercise. Check the beginning and end of the text and make sure that all unnecessary information like advertisements have been deleted. Then paste it into the blank field on the left and press “SEARCH” (Fig. 8).

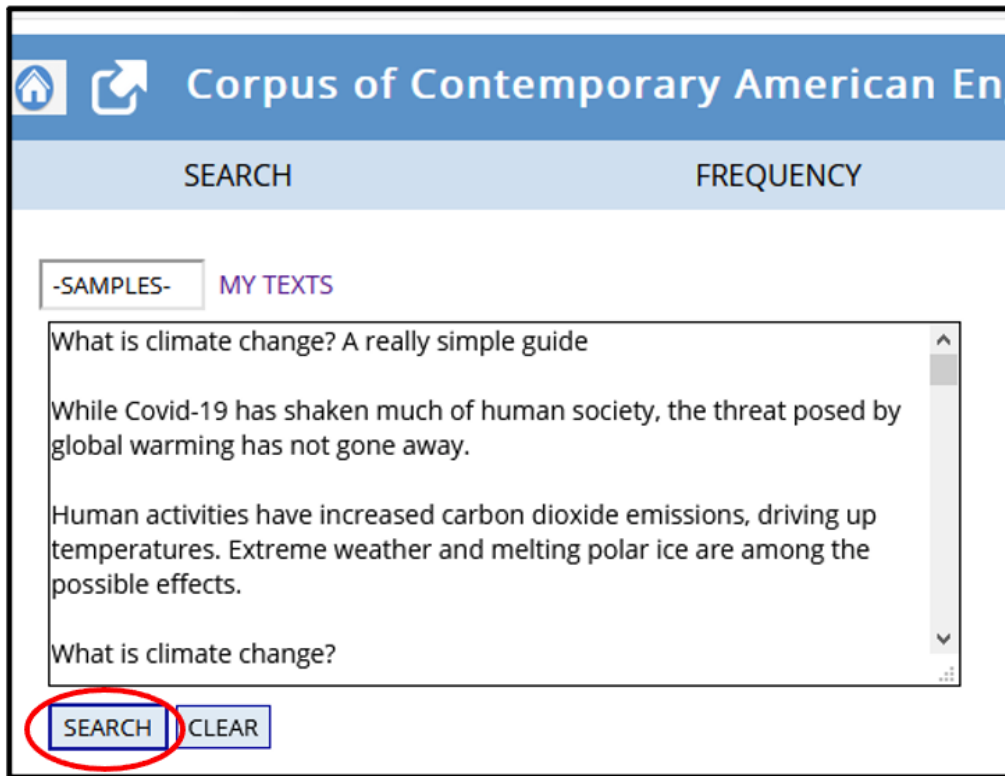


Fig. 8: Using the “Analyze text” function

You should now see the following page (Fig. 9).

EDIT TEXT	SAVE TEXT	<input checked="" type="radio"/> WORD	<input type="radio"/> PHRASE
FREQ RANGE	1-500	501-3000	> 3000
996 WORDS	49 %	16 %	20 %

CLICK ON ANY WORD BELOW FOR A FULL WORD SKETCH

What is **climate** change? A really **simple** **guide**

While Covid-19 has **shaken** much of human **society**, the **threat** **posed** by **global** **warming** has not gone away .

Human **activities** have increased **carbon dioxide emissions**, **driving** up **temperatures**. **Extreme** **weather** and **melting** **polar** **ice** are among the **possible** **effects** .

What is **climate** change ?

The Earth's **average** **temperature** is about 15C but has been much higher and lower in the past .

There are **natural** **fluctuations** in the **climate** but **scientists** say **temperatures** are now **rising** **faster** than at many other times .

(CLICK ANY WORD FOR FULL WORD SKETCH)

LOW FREQ	MID FREQ	HIGH FREQ
<p>9: climate</p> <p>6: atmosphere, greenhouse</p> <p>5: emissions</p> <p>4: carbon</p> <p>3: dioxide, melting</p> <p>2: dramatically, extreme, greenhouse-gas, methane, rapid, satellite</p> <p>1: absorbed, abundant, acidic, adapt, agriculture, alter, carbon-absorbing, complicated, concentration, coral, context-free, dramatic, extinctions, far-reaching, evaporates, exceed, flooding, floods, fluctuations, fossil, frequency, frozen, gateway, glaciers, globe, habitats, heatwaves, hostile, html, industrialisation, inland</p>	<p>9: warming</p> <p>6: global, temperature</p> <p>4: gases, released, scientists, temperatures</p> <p>3: energy, ice, likely, natural, rising, sea, surface</p> <p>2: activities, animals, average, century, characters, faster, gas, heats, hundreds, impact, mass, oceans, per, recent, rise, rises, society, storms, weather</p> <p>1: ability, according, affect, animal, argued, aspects, burned, burning, challenges, colder, contributing, current, dangerous, data, decades, decline</p>	<p>62: the</p> <p>28: and, of</p> <p>27: to</p> <p>24: in</p> <p>23: is</p> <p>14: are</p> <p>11: this</p> <p>10: change, more</p> <p>9: as, be</p> <p>8: a, by, will</p> <p>7: but, for, than</p> <p>6: from, it, much, that, water</p> <p>5: about, could, has, most</p> <p>4: also, at, can, changes, have, how, levels, number, there, what, years</p> <p>3: any, effect, effects, human, increase, increases, now, other, up, we, with, world, would</p> <p>2: all, areas, because, been</p>

Fig. 9: “Analyze text” – output

9. Save the text to be able to access it immediately in class without having to copy and paste it again. To do so, press “SAVE TEXT” (Fig. 10). In class, you can reuse it by simply clicking on “MY TEXTS” and the green tick (Fig. 11).

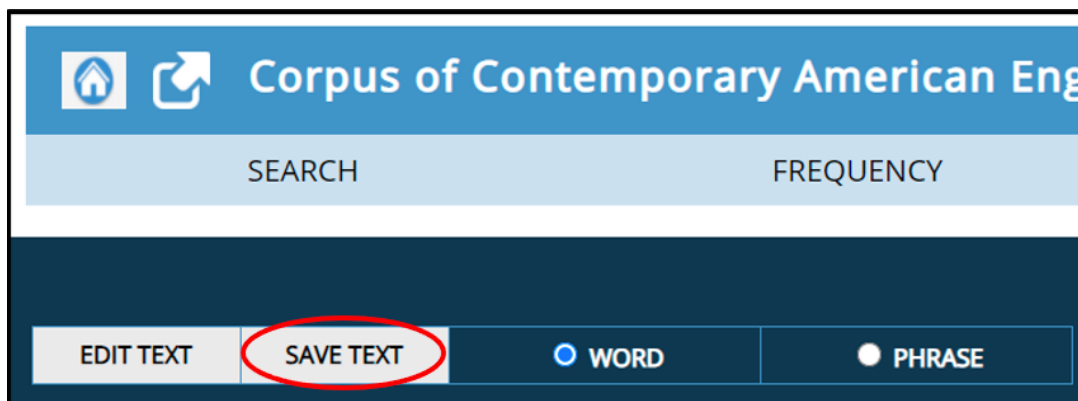


Fig. 10: Saving the analysed text

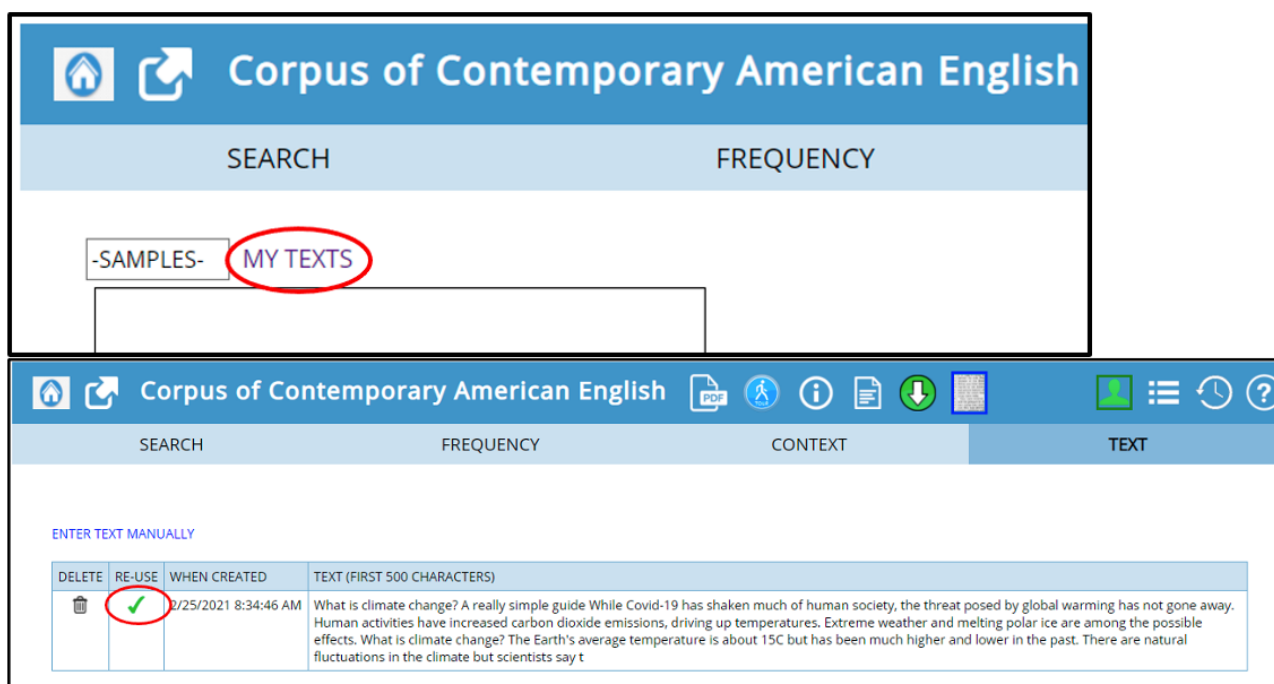


Fig. 11: Accessing saved texts

10. You will then be taken back to the page shown in Fig. 9, which will serve as the starting point for the first task of the lesson (see section 3.2.2).

3.1.2 Worksheet creation

Besides choosing and sharing an appropriate text, the teacher has to prepare a worksheet and an accompanying handout for the learners. Therefore, decide which specific vocabulary you want to focus on in advance.

In the Appendix, you will find a template for a worksheet and handout that guides the students through the exercises. They are tailored to the specific text mentioned above. However, feel free to use any other text you prefer. These worksheets may be used as a template to design your own. Both can be adapted to your own text and topic of choice. You will find the suggested solutions to the worksheet in the Appendix as well.

3.2 Lesson plan

3.2.1 Overview

Previous lesson: Introduction to new topic: climate change

Today's lesson: Learning words and collocations independently: Rising temperatures and climate science

Subsequent lessons: Presentations on climate change or writing a newspaper report

Table 1: Lesson plan

Phase	Task	Implementation	Interaction form	Medium
Warm-up (5 min)	Recapitulation of the last lesson on climate change. "Getting started" with english-corpora.org.	Class recapitulates the topic and its relevance. Teacher accesses english-corpora.org and goes to the "Analyze text" function.	Plenum	News article, images
Task 1 (10 min)	"Discovering new words".	Students learn how to look up unfamiliar words using the "Analyze Text" function; screenshare of teacher's COCA interface.	Plenum	News article, worksheet, computers, projector
Students' preparation (5 min)	"Getting started": starting computers, logging in on english-corpora.org	Preparation for following tasks	Individually	Worksheet, computer
Task 2 (20 min)	"Let's talk about temperature": finding out about collocations.	Learning vocabulary in chunks; students learn to use words in appropriate contexts.	Individually, partner work	Worksheet, handout, computers
Task 3 (optional) (20 min)	"Either... or" and "What can I say instead?"	Advanced students conduct more specific queries to acquire more detailed vocabulary knowledge.	Individually	Worksheet, handout, computers
Discussion (15 min)	Comparison and correction of the students' results.	Students present their solutions in class and reflect on the learned content and method.	Plenum	Projector/ Smart Board
Post-task/ homework (15 min)	"Word map": Organising new words and chunks.	Students integrate new vocabulary into already known structures.	Individually	Worksheet, handout, computer

3.2.2 Tasks

Task 1

Go to the "Analyse text" function of the COCA and access your saved text by as explained in section 3.1.2 above (steps 9 and 10).

From this point onwards, share your screen with the class and explain the following to your students:

The words in the text are highlighted in different colours. Each colour stands for a specific frequency range. Blue words are very frequent in English, green words occur less frequently. Yellow words are relatively rare. The latter are often very topic-specific and are therefore likely to be unfamiliar to students. Now, let the students begin their tasks. Ask them to read the text individually and to highlight unfamiliar words (task 1a). Let the students decide on three words they want to learn more about. On your shared screen, take the

students through the following steps for each word and explain what can be seen so that they can solve task 1b on their worksheets afterwards:

1. Make sure you are on the “WORD” interface (Fig. 12).

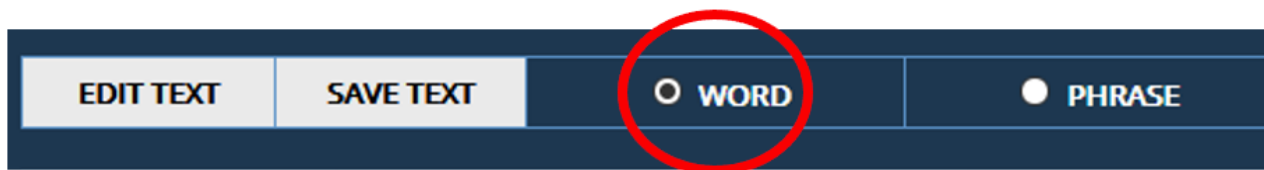


Fig. 12: The “WORD” interface

2. One by one, click on the three unfamiliar words in the text. A page with more information on the specific word will appear (see example in Fig. 13).

climate (NOUN) #1487

BLOG WEB TV/M SPOK FIC MAG NEWS ACAD

1. the weather in some location averaged over some long period of time 2. the prevailing psychological state **D M O C C G E**

YouGlish PlayPhrase Yarn
Translate: choose language

SYNONYMS (more)
atmosphere atmosphere, climate, environment, feeling, mood, sense, situation, surroundings weather environment, microclimate, temperature, weather

TOPICS (more)
greenhouse, warming, global, carbon, emission, temperature, drought, change, scientist, arctic, environmental, warm, dioxide, atmosphere, gas, tropical, fossil, energy, ocean, weather

COLLOCATES (more)
NOUN change, science, scientist, impact, model, effect, earth, policy
VERB change, affect, address, warm, predict, adapt, contribute, combat
ADJ global, political, current, economic, warm, cold, intergovernmental, changing
ADV eg, ie, negatively, radically, drastically, moderately, computationally, definitively

Fig. 13: Information page of the word climate

The main form of the word is highlighted and presented at the top of the page. It is followed by its word class in brackets. Below, the bar chart indicates the frequency distribution of the word across the eight text registers included in the COCA: blogs, websites, TV and movies, spoken language, fiction, magazines, news and academic publications. For example, the noun *climate* is least common in TV/movie language and fiction, and most common in blogs and academic language. The section “Topics” in the top-right corner shows words associated with *climate* and can help you find related words and concepts.

3. In order to fill in Table 1 on your worksheet, find the word’s definitions below the bar chart. Also, look at the section below that shows synonyms and marks the most appropriate ones, as well as the section “Collocates”, listed on the right.

4. To find out more about the word’s collocates, click on “more” right next to it (Fig. 13). You will then be taken to the following page (Fig. 14).

COLLOCATES **CLIMATE** **NOUN** Advanced options [Collocates](#) [Clusters](#) [Topics](#) [Dictionary](#) [Texts](#) [KWIC](#) [↻](#) [↓](#)

+ NOUN	NEW WORD	?	+ ADJ	NEW WORD	?	+ VERB	NEW WORD	?	+ ADV	NEW WORD	?
23780	7.29	change	2455	5.52	global	1602	2.94	change	93	5.98	eg
1925	4.39	science	1476	3.32	political	656	3.59	affect	40	3.59	ie
1805	5.29	scientist	1048	3.73	current	492	3.29	address	20	2.58	negatively
1352	4.59	impact	761	3.07	economic	382	5.01	warm	19	2.52	radically
1306	3.66	model	741	4.40	warm	276	3.35	predict	16	2.65	drastically
1073	3.20	effect	560	3.28	cold	242	4.18	adapt	12	2.51	moderately
796	3.36	earth	488	9.37	intergovernmental	216	2.61	contribute	8	5.70	computationally

Fig. 14: "Collocates" page of the word *climate*

There, you will see four tables grouped by part of speech and frequency, one each for noun, adjective, verb and adverb collocates. Open the "Advanced Options" menu, enter the number 1 into the blank field and press GO (Fig. 15). Now, only collocates appearing to the immediate left or right of the noun *climate* will be displayed in the tables.

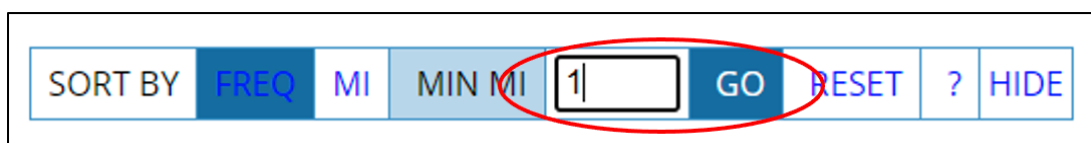
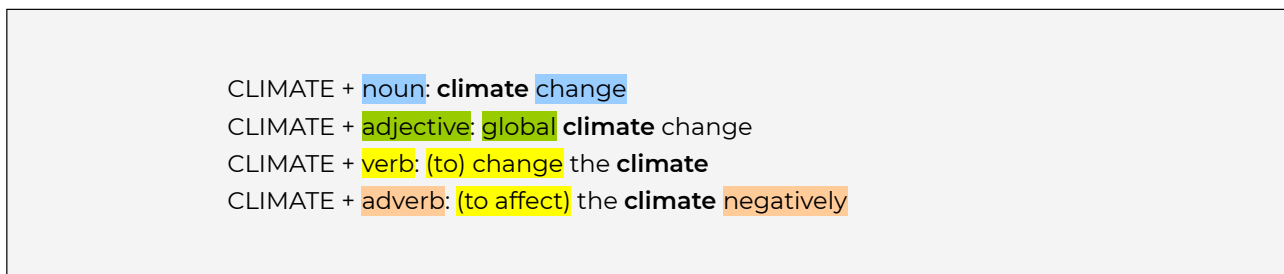


Fig. 15: "Advanced options" menu on the "Collocates" page

5. Pay attention to the grey blocks in the columns next to each collocate (Fig. 14). They indicate the most frequent position of the collocate relative to the search word *climate*, i.e., either before or after the word. The darker the colour, the more frequently does the search word's collocate appear in this position. For example:



In some cases, the blocks to the left and right of the collocate are white, which means that these collocates can appear both before and after the noun *climate*.

6. To go back to the search word's information page, either click on the little arrow in the top-left corner or on "ANALYZE TEXT" (Fig. 16).

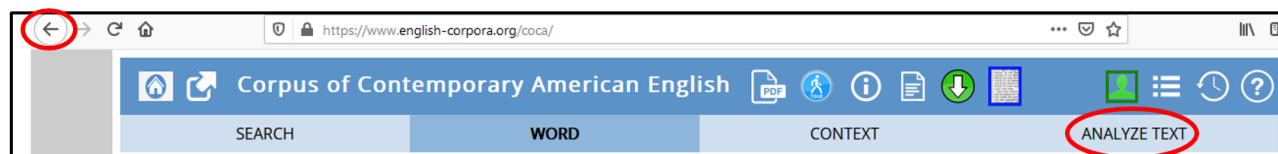


Fig. 16: Going back to the information page

Task 2 and 3

After guiding the students through the first task, ask them to start their computers and to solve tasks 2 and 3 independently with the help of the handout (cf. Appendix).

4 Follow-up activities and further ideas

As shown in the worksheet and the lesson plan above, this lesson can be adapted to different levels of language proficiency and the amount of experience students have had with using digital media in the classroom. Task 3 could be given as homework or as an additional task for more advanced students, for example.

Further, this lesson can serve as a basis for different follow-up activities. Apart from finishing and expanding the word map on the worksheet, further activities may include:

- gathering new words relating to the “Fridays for Future” movement,
- preparing presentations about climate-related topics,
- working in groups and designing posters to inform others about climate-related issues or movements.

5 Conclusion

This lesson allows EFL learners to expand their vocabulary while working independently. With the help of a corpus, students develop a deeper understanding of topic-specific vocabulary. The worksheet aims to show students how to learn new vocabulary in chunks, which then enables them to use new words in appropriate contexts. Organising new vocabulary in a word map makes learning more sustainable, as visualisations can help learners memorise new words and collocations more easily.

Students are provided with a detailed handout that guides them through the individual steps of the lesson so that an introductory lesson on how to use a corpus is not needed.

Crucially, the methods used in this lesson can easily be applied to other topics. To do so, you would only need to change the basic text and adapt the worksheet and handout.

6 Resources and references

Resources:

Corpus of Contemporary American English (COCA). <https://www.english-corpora.org/coca/>.

Pomikalek, Jan. (2021). JusText demo. <https://nlp.fi.muni.cz/projects/justext/>.

References:

BBC News (18 November 2020). “What is climate change? A really simple guide”. <https://www.bbc.com/news/science-environment-24021772>. (25 March, 2021).

Conrad, Susan. 2016. Using corpus linguistics to improve the teaching of grammar. In Eli Hinkel (ed.), *Teaching English grammar to speakers of other languages*, 38-62. Florence: Taylor and Francis.

Friginal, Eric. 2018. *Corpus linguistics for English teachers: new tools, online resources, and classroom activities*. New York, NY: Routledge.

Gill, Mehr. (20 September, 2019). Explained: What is the Global Climate Strike movement? *The Indian EXPRESS*. <https://indianexpress.com/article/explained/explained-what-is-the-global-climate-strike-movement-6013302/> (24 March, 2021).

Niedersächsisches Kultusministerium. 2015. Kerncurriculum für das Gymnasium, Schuljahrgänge 5-10: Englisch. https://cuvo.nibis.de/cuvo.php?skey_lev0_0=Schulbereich&svalue_lev0_0=Sek+I&skey_lev0_1=Fach&svalue_lev0_1=Englisch&skey_lev0_2=Dokumentenart&svalue_lev0_2=Kerncurriculum&docid=1197&p=detail_view (25 March, 2021).

Niedersächsisches Kultusministerium. 2017. Kerncurriculum für das Gymnasium – gymnasiale Oberstufe, die Gesamtschule – gymnasiale Oberstufe, das Berufliche Gymnasium, das Abendgymnasium, das Kolleg: Englisch. https://cuvo.nibis.de/cuvo.php?skey_lev0_0=Dokumentenart&svalue_lev0_0=Kerncurriculum&skey_lev0_1=Fach&svalue_lev0_1=Englisch&docid=1132&p=detail_view (25 March, 2021).

7 Appendix

Download the full worksheet and the handout by clicking on the images.

Student worksheet: Climate change – new words

How to solve these tasks successfully:

Check if the following things are right. If you cannot tick off all of them, tell your teacher about the problem. As soon as everything is in order, you are ready to start!

- My computer is running.
- It is connected to the internet.
- I have the text my teacher provided ready for digital use.
- I received a handout.



Tasks

1) Discovering new words

- a. Read the article and highlight unfamiliar words.
- b. Pick three unfamiliar words you want to learn more about. Fill in Table 1 below.

Table 1: Information about new words (N = noun, V = verb, ADJ = adjective, ADV = adverb)

Unfamiliar word	Synonyms	Definition	Most frequent collocates
Example: <i>climate</i>	<i>atmosphere, weather</i>	<i>The weather in some location averaged over some long period of time</i>	1: <i>climate change</i> 2: <i>global climate</i> 3: <i>climate science</i>
...			1: 2: 3:
...			1: 2: 3:
...			1: 2: 3:

Fig. 17: Student worksheet

Student handout: Climate change – new words**Task 2)**

a. What kinds of temperatures and temperature phenomena exist?

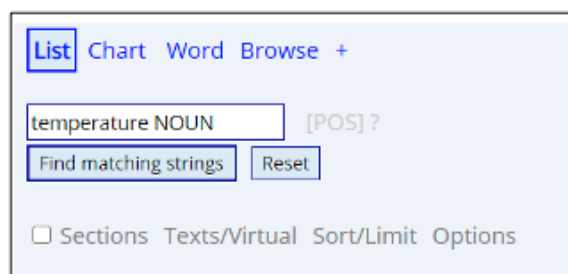
1. Go to www.english-corpora.org/ and choose the Corpus of Contemporary American English (COCA).



The screenshot shows the English-Corpora.org website. The header includes the logo and navigation links: home, corpora, users, related resources, my account, upgrade, help. Below the header, there is a description of the site as the most widely used online corpora, with links for a guided tour and search types. A table lists various corpora with columns for Corpus (online access), Download, # words, Dialect, Time period, and Genre(s).

Corpus (online access)	Download	# words	Dialect	Time period	Genre(s)
Web: The Intelligent Web-based Corpus		14 billion	6 countries	2017	Web
News on the Web (NOW)		11.9 billion+	20 countries	2010-yesterday	Web: News
Global Web Based English (GloWBE)		1.9 billion	20 countries	2012-13	Web (incl blogs)
Wikipedia Corpus		1.9 billion	(Various)	2014	Wikipedia
Corpus of Contemporary American English (COCA)		1.0 billion	American	1990-2019	Balanced
Coronavirus Corpus		844 million+	20 countries	Jan 2020-yesterday	Web: News
Corpus of Historical American English (COHA)		400 million	American	1840-2000	Balanced

2. On the COCA search interface, type in the word “temperature” followed by “NOUN”.



The screenshot shows the search interface for the COCA corpus. It includes a search bar with the text "temperature NOUN" and a "[POS]?" label. Below the search bar are buttons for "Find matching strings" and "Reset". At the bottom, there is a checkbox for "Sections Texts/Virtual Sort/Limit Options".

3. Click on “Options” and change “GROUP BY WORDS” to “GROUP BY LEMMAS” in the drop-down menu.

Fig. 18: Student handout

12. Act responsible, think sustainable – Expanding vocabulary to write about sustainability

RIEKE STAHNKE, MARLENA WEGESIN, AND LARA BREDE

1 Introduction and rationale

Project-based teaching has become increasingly popular. The method is deemed to be potentially highly beneficial in terms of student engagement, independent and social learning, as well as offering many options for differentiation within mixed ability groups. However, conducting projects is often labour-intensive for teachers. Especially in English Language Teaching (ELT), project work often requires students to acquire subject-specific linguistic skills.

This chapter aims to present a flexible approach for providing students with the necessary linguistic support to conduct project work while also keeping the time needed for teacher preparation at a minimum. By teaching students how to use basic strategies of corpus linguistic research, this method will enable students to find the lexical information they need for the specific topic of their choice. This allows for more freedom and differentiation, and also guides the students to think of language as lexical patterns rather than as strings of individual words.

As an exemplary context in which the presented approach can be taken, this chapter outlines a project about *sustainability*, which is a highly topical issue in today's world and, thus, very relevant to students' lives. The main objective is to enable students to talk and write about topics of their own interests related to *sustainability* and to express personal opinions and emotions about them, which is in line with the secondary school curriculum of Lower Saxony, Germany (Niedersächsisches Kultusministerium 2015: 19, 23).

Outline and learning objectives

- Target learners: Intermediate to upper-level students at secondary school
- Intended product of the project: A blog entry, podcast or video
- Prior knowledge: Students have acquired some knowledge and key vocabulary on *sustainability*
- Previous lessons: In groups, students have decided on their final topics for their projects and created a draft outline for their writing including key topics and arguments

Preparatory lesson 1 introduces data-driven learning using the “List”-function

- Preparation time: 20 min (or less if you are already familiar with www.english-corpora.org)

- Lesson time: 45 min

Practice lesson 2 uses data-driven learning to find collocations for key vocabulary for the students' project ideas

- Preparation time: 25 min
- Lesson time: 90 min

Learning objectives

By the end of the second lesson, students should...

- understand the benefits of using corpora to express themselves in idiomatic and natural English,
- understand the meaning and use of collocations,
- be able to use basic functions of the corpus tool to independently search for idiomatic language (word pairs, collocations and example sentences) relevant to their project topic, and
- have contextualised and expanded their existing vocabulary about sustainability in preparation for working on their projects.

2 Corpus, tools and methods

The corpus used in this chapter is the Corpus of Contemporary American English (COCA), freely accessible on www.english.corpora.org/COCA. It provides the option to look exclusively at the language of blogs (assuming that the language of blogs is comparable to the language of podcasts and videos), and includes a large number of recent texts, which is vital for a topic like sustainability that has rapidly gained public attention in the past few years. Moreover, the COCA works best for finding collocations on www.english-corpora.org, as it automatically groups words by part-of-speech, making its output more learner-friendly.

In the lessons, students will first use the “Texts/Virtual”-function to create a virtual corpus (their own subcorpus of the COCA) that only includes blogs about sustainability. They will then continue to explore this subcorpus by means of the “List”-, and “Collocates”-functions to find idiomatic language.

3 Step-by-step guide


3.1 Teacher Preparation

Creating a virtual corpus

Go to www.english-corpora.org. You should see a list of the corpora available on the website. Click on “Corpus of Contemporary American English (COCA)” (Fig. 1):



The most widely used online corpora. [Guided tour](#), [overview](#), [search types](#), [variation](#), [virtual corpora](#), [corpus-based resources](#).

The links below are for the online interface. But you can also  download the corpora for use on your own computer.

Corpus (online access)	Download	# words	Dialect	Time period	Genre(s)
iWeb: The Intelligent Web-based Corpus		14 billion	6 countries	2017	Web
News on the Web (NOW)		11.8 billion+	20 countries	2010-yesterday	Web: News
Global Web-Based English (GloWbE)		1.9 billion	20 countries	2012-13	Web (incl blogs)
Wikipedia Corpus		1.9 billion	(Various)	2014	Wikipedia
Corpus of Contemporary American English (COCA)		1.0 billion	American	1990-2019	Balanced
Coronavirus Corpus		830 million+	20 countries	Jan 2020-yesterday	Web: News
Corpus of Historical American English (COHA)		400 million	American	1810-2009	Balanced
The TV Corpus		325 million	6 countries	1950-2018	TV shows
The Movie Corpus		200 million	6 countries	1930-2018	Movies
Corpus of American Soap Operas		100 million	American	2001-2012	TV shows

Fig. 1: List of corpora available on the web interface

Make sure that you are logged in, as you will create a virtual corpus on your account, i.e., it will only be visible to you (Fig. 2) (If you have not used the website before, you will need to register first by clicking on “My account” in the menu and then “Register” and filling out the required form. This is free and takes no longer than a minute).

Select “Texts/Virtual” in the blue box on the right, and click on “Create corpus “(Fig. 2).

Fig. 2: Creating a virtual corpus

You should now see a search form (Fig. 3) that allows you to select certain texts from the COCA to include in your virtual corpus. For our purposes, we only want to include texts of the genre “Blog”. However, to exclude highly specialized texts with academic language, we will not select the subgenres “ACAD” and “LEGAL”. To do so, press “Ctrl” on your keyboard and click the individual subgenres of “BLOG” in the list, leaving out “ACAD” and “LEGAL” (Fig. 3).

To only include texts that are relevant to the topic of our project, type *sustainability* in the field next to “Words in text”. Then press “Submit” (Fig. 3).

Corpus of Contemporary American English

SEARCH VIRTUAL CORPUS CONTEXT OVERVIEW

Source
(optional, and can use substring)

Article title Include plots (TV and Movies)

Years -

Genre/domain

WEB	BLOG	MOV	TV	SPOK	FIC	MAG	NEWS	ACAD
ALL Acad Argum Fic Info Instr	Acad Argum Fic Info Instr Legal	ALL Action Adult Advntr Anim Biog	ALL Action Advntr Anim Comedy Crime	ALL ABC NBC CBS CNN FOX	ALL Gen (Book) Gen (Jrnl) SciFi/Fant Juvenile Movies	ALL News/Opin Financial Sci/Tech Soc/Arts Religion	ALL Money Life Sports Editorial Misc	ALL Education History Geog/SocSci Law/PolSci Humanities

Words in text

You can create a "virtual corpus" by selecting texts from among the nearly [500,000+ texts](#) in the COCA corpus. This allows you to move beyond the categories that we have created (by genre and year) to select texts from the corpus in any way that you want.

You can use any of the criteria shown to the left to determine what texts will be in the virtual corpus. All of the criteria are optional – just use those that you want.

Note that via this search form, you can only enter one single word in the [WORD] field. If you want to do more complicated searches with multi-word strings, substrings, lemmas, grammatical constructions, etc then you should create a virtual corpus using keywords. See the instructions by clicking on TEXTS/VIRTUAL in the search form, and then "Keywords in the Text" (the first link in that help file).

After you've created the corpus, you can add or delete texts and make other modifications to the corpus. You can also search just within the virtual corpus that you've created; you can compare the frequency of words and phrases across your different virtual corpora; and you can generate keyword lists from any of the virtual corpora.

Fig. 3: Searching for texts in the corpus

On the next page, you will find a list of all the texts your query has returned. Give your virtual corpus a name (e.g., *sustainability*) by typing it in the field next to “SAVE AS” and finally click “SUBMIT” (Fig. 4) to finish building your virtual corpus (Fig. 5).

Corpus of Contemporary American English

SEARCH VIRTUAL CORPUS CONTEXT HELP

You can add or remove texts below, and then either save these texts as a new virtual corpus, or else add them to an existing virtual corpus.

SAVE AS: Sustainability Blog OR ADD TO: --SELECT-- SUBMIT RESET

HELP	<input type="checkbox"/> 100	YEAR	GENRE	SOURCE	TITLE
1	<input checked="" type="checkbox"/>	2012	BLOG	...nationaljournal.com	Should Fuel Taxes Pay For Alternative Transportation...
2	<input checked="" type="checkbox"/>	2012	BLOG	...nationaljournal.com	What Are The Five Most Important Issues Facing The Aviation Industry...
3	<input checked="" type="checkbox"/>	2012	BLOG	...nationaljournal.com	What Do You Think Of DOT's Draft Strategic Plan? - Transportation...
4	<input checked="" type="checkbox"/>	2012	BLOG	...a-wildwood.patch.com	County 'Goes Ape' to Help Save Greensfelder Park in Wildwood...
5	<input checked="" type="checkbox"/>	2012	BLOG	...al.michaelbunker.com	IndieGoGo Stole the Money... (as far as we know now)...
6	<input checked="" type="checkbox"/>	2012	BLOG	...alismaccelerator.com	The Role of Philanthropy in Local Accountability Journalism...
7	<input checked="" type="checkbox"/>	2012	BLOG	...allupblog.gallup.com	Gallup.com - The Gallup Blog: A Different Way to Think About...
8	<input checked="" type="checkbox"/>	2012	BLOG	...amethepatriarchy.com	Spinster aunt is informed that Pussy Riot's ideological carpet doesn't...
9	<input checked="" type="checkbox"/>	2012	BLOG	...ance.fortune.cnn.com	Congratulations, Mr. President. Here's how to fix the economy. - The...
10	<input checked="" type="checkbox"/>	2012	BLOG	...anover-workedmom.com	Help Educate Children about the Importance of Healthy Diets and...
11	<input checked="" type="checkbox"/>	2012	BLOG	...antafe.wordpress.com	A Little Help All Around A Year in Santa Fe...
12	<input checked="" type="checkbox"/>	2012	BLOG	...asedal.wordpress.com	Anyone over the age of 35 should read this, as I copied this from a...
13	<input checked="" type="checkbox"/>	2012	BLOG	...atsforlunchhoney.net	What's For Lunch Honey? Experience Your Senses: Dubai: A State...
14	<input checked="" type="checkbox"/>	2012	BLOG	...aur-net.blogspot.com	Pterosaur.net Blog: How giant pterosaurs are struggling to take off...
15	<input checked="" type="checkbox"/>	2012	BLOG	...bcentre.blogspot.com	GEO REHAB CENTRE: USE OF MASS MEDIA IN THE AREA OF...
16	<input checked="" type="checkbox"/>	2012	BLOG	...bergin.wordpress.com	Is Sustainability an important part of a design brand? rory bergin's...
17	<input checked="" type="checkbox"/>	2012	BLOG	...bility.wordpress.com	The power of paradigms in land management: the case of grazing...
18	<input checked="" type="checkbox"/>	2012	BLOG	...blogus.wordpress.com	Leaves and Twigs: An Unscientific Roundup of the Best...

Fig. 4: List of texts selected

Corpus of Contemporary American English

SEARCH VIRTUAL CORPORA CONTEXT HELP

MY VIRTUAL CORPORA

HELP	↑ ↓	LIST NAME ↓	# TEXTS ↓	# WORDS ↓	FIND KEYWORDS ● SPECIFIC ○ FREQ	CREATED ↓
1		SUSTAINABILITY BLOG	744	3,624,790	NOUN VERB ADJ ADV N+N ADJ+N	0 h

Fig. 5: Your virtual corpus

Preparing the warm-up activity

The warm-up activity aims to show students the benefits of using corpora to obtain information about how words are naturally used together by asking what verb (*maintain*, *preserve*, *retain*, etc.) is most frequently associated with *biodiversity*. As preparation, search for the results in the full corpus:

Choose the COCA, click on the “+” next to “Browse” and then on “Collocates”.

Then type *biodiversity* into the “Word/phrase” search field, click on “POS” (part of speech) next to “Collocates”, choose “verb.ALL”, and select one position to the left (= verb that precedes *biodiversity* by one position within the sentence).

Next, click on “Options”, and let your search results be grouped by “LEMMAS” to make sure that conjugated verb forms do not result in duplicated findings (Fig. 6).

The screenshot shows the 'Collocates' search interface. At the top, there are navigation tabs: 'List', 'Chart', 'Word', 'Browse', 'Collocates' (highlighted), 'Compare', and 'KWIC'. Below these, there is a search area with a text input containing 'biodiversity' and a label 'Word/phrase [POS]'. To the right, there is a 'Collocates' dropdown menu set to 'verb.ALL'. Below the search area is a frequency bar with buttons for '+', '4', '3', '2', '1', '0', '0', '1', '2', '3', '4', and '+'. The '1' button is highlighted in green. Below the frequency bar are two buttons: 'Find collocates' and 'Reset'. At the bottom, there is an 'Options' section with a checkbox and the text 'Sections Texts/Virtual Sort/Limit Options'. The 'Options' section is expanded, showing several settings: '# HITS' set to '100', '# KWIC' set to '200', 'GROUP BY' set to 'LEMMAS', 'DISPLAY' set to 'RAW FREQ', and 'SAVE LISTS' set to 'NO'.

Fig. 6: Searching for verb collocates of biodiversity

In this case, do not choose the newly-created virtual corpus, as the whole content of the COCA will provide more results.

Submit your search (“Find collocates”) to obtain a list of the most frequent verbs co-occurring with *biodiversity* (Fig. 7).

ON CLICK: [CONTEXT](#) [TRANSLATE \(??\)](#) [GOOGLE](#) [IMAGE](#) [PRON/VIDEO](#) [BOOK](#) (HELP)

HELP			FREQ
1	<input type="checkbox"/>	[CONSERVE]	83
2	<input type="checkbox"/>	[PROTECT]	74
3	<input type="checkbox"/>	[PRESERVE]	29
4	<input type="checkbox"/>	[REDUCE]	28
5	<input type="checkbox"/>	[PROMOTE]	25
6	<input type="checkbox"/>	[MAINTAIN]	23
7	<input type="checkbox"/>	[SEE]	17
8	<input type="checkbox"/>	[BE]	16
9	<input type="checkbox"/>	[THREATEN]	14
10	<input type="checkbox"/>	[AFFECT]	13
11	<input type="checkbox"/>	[ENHANCE]	10
12	<input type="checkbox"/>	[ALTER]	9
13	<input type="checkbox"/>	[SAVE]	8
14	<input type="checkbox"/>	[INCREASE]	7
15	<input type="checkbox"/>	[DESTROY]	7
16	<input type="checkbox"/>	[SUSTAIN]	7
17	<input type="checkbox"/>	[CONSIDER]	6
18	<input type="checkbox"/>	[INTEGRATE]	6
19	<input type="checkbox"/>	[INVESTIGATE]	6
20	<input type="checkbox"/>	[INFLUENCE]	6

Fig. 7: List of verb collocates of the search item biodiversity

Take a screenshot which you can present to your students during the warm-up of the second lesson.

For the first warm-up activity, select the first five findings, and click on “CONTEXT” just above the first verb of the list (Fig. 7).

Randomise your results afterwards (by clicking on a number next to “Find sample” in the top left corner of the result list) and take a screenshot to present in class.

Getting to know the “List”-function

Next, you should become familiar with the tasks that are used in the lesson and try out the “List”-function so that you can walk your students through the steps in class.

Go to the COCA again and type the word *sustainability* in the search box. This word serves as an example to show how the “List”-function works.

Limit your search to your virtual corpus by clicking on “Text/Virtual”, select your corpus (here: *sustainability*) and click on “Find matching strings” (Fig. 8).



Fig. 8: Using the “List”-function

You will find a very short table that tells you how often *sustainability* appears in the corpus. Click on the word itself (Fig. 9) to obtain a list of examples of texts that feature the noun *sustainability* (Fig. 10).

Corpus of Contemporary American English			
SEARCH	FREQUENCY	CONTEXT	OVERVIEW
ON CLICK: CONTEXT TRANSLATE (??) GOOGLE IMAGE PRON/VIDEO BOOK (HELP)			
HELP	?	ALL FORMS (SAMPLE): 100 200 500	FREQ
1	<input type="checkbox"/>	SUSTAINABILITY	1332
			0.494 seconds

Fig. 9: Frequency results of sustainability

These examples are called **concordance lines**. Click on any number in the top left corner (Fig. 10) to randomise the search and thus gain a more varied overview of how the word is used in different contexts.

Corpus of Contemporary American English

SEARCH FREQUENCY CONTEXT OVERVIEW

FIND SAMPLE: 100 200 500 1000
PAGE: << < 1 / 14 > >>

CLICK FOR MORE CONTEXT [?] SAVE LIST CHOOSE LIST CREATE NEW LIST [?]

Rank	Year	Type	Source	Context
1	2012	BLOG	...s.kontentkonsult.com	A B C Instead of economics being our tool for societal advancement, happiness, health, and sustainability , the profession has lost its way in econometrics, neo-
2	2012	BLOG	100rsns.blogspot.com	A B C know that tenured faculty aren't about to take a pay cut to ensure the sustainability of the system. # I understand, but that's not an excuse.
3	2012	BLOG	cruxcatalyst.com	A B C and they might even be figments of our collective imagination. # Sharon is a sustainability ideas transmitter, writer and activist, from Adelaide, South Aus
4	2012	BLOG	cruxcatalyst.com	A B C from 'bigger' to 'better'. # Sharon has been working on sustainability issues in paid and voluntary work in both the government and the NGO sector since
5	2012	BLOG	dailykos.com	A B C will in fact actively oppose it? # It seems to me that the sustainability and potential scale of a "Clintonian-German" model remains to be seen. The
6	2012	BLOG	dailykos.com	A B C is that unless it is addressed in the context of environmental issues, the economic sustainability and soundness of alternative contexts will become irrele
7	2012	BLOG	blog.icfull.com	A B C in the electronic information sector collective uplift of the circumstances, we believe that the sustainability of investment opportunities in the industry sh
8	2012	BLOG	urbnlivn.com	A B C Future # Tomorrow, a two-story house that was designed to be a model for sustainability will debut at the Seattle Center. Dubbed the 'House of the imm
9	2012	BLOG	www3.gmiratings.com	A B C but significant positive relationship with the level of relative stock prices, as do corporate sustainability efforts. And, of course, there's the catch-all of all i
10	2012	BLOG	lifetips.att.com	A B C and I'm done! Meet the Author # I became Director of the Sustainability Program Office, AT &T; Marketing Communications in April, 2011. My AT &T;
11	2012	BLOG	lifetips.att.com	A B C and a background in Marketing, I assist AT &T's: corporate customers in meeting their sustainability goals. I am also an author on AT &T's: Business Netw
12	2012	BLOG	livinggreenmag.com	A B C . is your guide to 'the mindful life': yoga, organics, sustainability , genuine spirituality, conscious consumerism, fair fashion, the contemplative arts... anyth
13	2012	BLOG	...e-caeli.blogspot.com	A B C 3203956 Suppliers Sustainability consultants Is Sustainability an important part of a design brand? # Peter Murray
14	2012	BLOG	...e-caeli.blogspot.com	A B C 3203956 Suppliers Sustainability consultants Is Sustainability an important part of a design brand? # Peter Murray, chairman of the
15	2012	BLOG	...bergin.wordpress.com	A B C for designers he said that it no longer made sense for consultancies to talk about sustainability being part of their brand, as it' ought to be part of every
16	2012	BLOG	...bergin.wordpress.com	A B C that he is partly right and partly wrong. He ought to be right because sustainability ought to be part of how people work in construction by now, we certa
17	2012	BLOG	...bergin.wordpress.com	A B C Or just talking about it. # I don't think that we are doing sustainability yet. We are still routinely designing and building buildings that fall very far short.
18	2012	BLOG	...bergin.wordpress.com	A B C mood in the housing industry, still reeling from the credit crunch, is that sustainability can wait until we can afford it. So how can we claim to be
19	2012	BLOG	...bergin.wordpress.com	A B C that we should accept lightly. # I think that designers need to think about sustainability more, and what it means to us both as individuals and practices,

Fig. 10: Search result using the "List"-function

3.2 Lesson 1

Plan for the first lesson

Duration: 45 minutes

Table 1: Lesson plan for lesson 1

Phase	Task	Implementation	Interaction form	Materials
(1) Warm-Up 5 min	Translating: Wir müssen die Artenvielfalt erhalten.	Problematisation of translating <i>etwas erhalten</i> without considering the context	Plenum	Electronic devices*
(2) Preparation 15 min	Preparation of corpus work	Creating an account and a virtual corpus on www.english-corpora.org Following the instructions on the worksheet	Individual work	Electronic devices, Worksheet A
(3) Practice 15-20 min	How can we use <i>sustainable</i> in a sentence?	Using the "List"-function and working with concordance lines Practising data-driven learning	Individual/ partner work	Electronic devices, Worksheet A
(4) Conclusion 5-10 min	How did it go?	Sharing results and discussing problems Outlook on next lesson	Plenum	---

* = interactive whiteboard, students' computers/laptops

Tasks and instructions

Warm-up

Wir müssen die Artenvielfalt erhalten.

The warm-up activity serves not only to activate students' cognitive skills, but also to make them aware of the relevance of corpus work for being able to express themselves accurately and idiomatically in English. For this purpose, select a topic-related sentence in the students' first language which includes at least one word that is problematic with regard to its translation into English. In this chapter, the German sentence *Wir müssen die Artenvielfalt erhalten* ('We need to conserve biodiversity') has been chosen to highlight the ambiguous meaning of the verb *erhalten*, which, in other contexts, could just as well be translated as *keep*, *retain* or *get*.

Present the sentence on the board and ask your learners to translate the sentence into English. It is assumed that students will be divided over the translation of the verb *erhalten* and you can show them the respective entry in a learner dictionary (Fig. 11) to make them aware of the different options for communicating the meanings of *erhalten*.

The image shows a screenshot of a learner dictionary entry for the German verb 'erhalten'. It is divided into two main sections: '4. erhalten (bewahren):' and '5. erhalten BAU (bewahren):'. Each section lists various grammatical uses of the verb with their corresponding English translations. Each entry includes a speaker icon and a three-dot menu icon. At the bottom of each section, there is a link to '32 Beispiele aus dem Internet' for section 4 and '21 Beispiele aus dem Internet' for section 5.

4. erhalten (bewahren):		
etw erhalten	to maintain sth	
etw erhalten (Vitamine/Wirkstoffe)	to retain	
[durch etw Akk] erhalten bleiben	to be preserved [by sth]	
[sich Dat] etw erhalten	to keep sth	
ich sehe, du hast dir deinen Optimismus erhalten	I see you're still an optimist	
32 Beispiele aus dem Internet		
5. erhalten BAU (bewahren):		
etw erhalten	to preserve sth	
etw ist erhalten	sth is preserved	
gut erhalten sein scherzh ugs	to be well-preserved scherzh ugs	
[durch etw Akk] erhalten bleiben	to remain preserved [by means of sth]	
jdm erhalten bleiben	to be with sb	
jdm erhalten bleiben iron	to be with sb	
jdm erhalten bleiben iron	to not lose sb	
21 Beispiele aus dem Internet		

Fig. 11: Dictionary entry for *erhalten* (www.pons.de)

Seeing that there are obviously several possible English translations (*maintain, retain, keep, preserve*), students will understand that they need another tool to find out which of the translations is appropriate in the example sentence. This point gives you the perfect opportunity to introduce the world of corpora and corpus work generally by subsequently presenting the concordance lines for “verb + *biodiversity*” using the “Collocates”- function on www.english-corpora.org (Fig. 12).

The screenshot shows the COCA interface with the search results for 'conserve biodiversity'. The interface includes a navigation bar with 'SEARCH', 'ERROR', 'CONTEXT', and 'CONTEXT +'. Below the navigation bar, there are search controls and a table of results. The table has columns for line number, year, source, and the concordance line itself. The concordance lines are highlighted with green boxes around the verb 'conserve biodiversity'.

Line	Year	Source	Concordance Line	
1	1997	ACAD	EnvirAffairs	should awaken a broader cross-section of the world's peoples to the need to conserve biodiversity in order to maintain sustainability of the Earth's vanishi
2	1991	MAG	Sierra	a time, of even a single national park. # The movement to preserve biodiversity represents a natural evolution of conservation strategy. It is both the next
3	2004	ACAD	Environment	, 2003). # DIAGRAM: Figure 1. New market solutions to conserve biodiversity : NOTE: OTC (" over-the-counter ") trading involves direct negotiation with buy
4	2010	SPOK	NPR_Science	, to do is think in a much more complicated way about how to preserve biodiversity . We know that simply putting a fence around a protected area and lea
5	1996	MAG	Environmental	think we have to come up with a formula that will enable us to preserve biodiversity without necessarily managing it in the sense of trying to freeze it in tin
6	1997	ACAD	Bioscience	# Virtually all conservation biologists agree that habitat preservation is the best way to conserve biodiversity . Setting aside large tracts of land so that they
7	2003	MAG	NaturalHist	nation, the demands of economic development must compete for land with efforts to conserve biodiversity . In that context, the study of endemism can be
8	2004	ACAD	Bioscience	or the only answer. Protected areas, although necessary, are insufficient to conserve biodiversity in the long term (Soul and Terborgh 1999). Some species,
9	2009	MAG	MotherJones	forests: What happens to the people who use the land? Efforts to protect biodiversity in the dwindling wildlands of the world have increasingly run into a d
10	1997	ACAD	EnvirAffairs	which is implementation by individual countries of new national laws and policies designed to conserve biodiversity resources and to use those resources
11	2000	ACAD	Humanist	the United Kingdom and elsewhere spearheaded the fledgling international program to protect wildlife and preserve biodiversity . The American Humanist
12	2012	WEB	rassak.com	of YouTube from the Rainforest Alliance, a group with a mission to " conserve biodiversity and ensure sustainable livelihoods by transforming land-use pra
13	1992	ACAD	Bioscience	21 " as drafted for Rio also calls for a variety of actions to conserve biodiversity (UNEP 1992b). # Indeed, the United States has been conserving genetic
14	2000	ACAD	Bioscience	, biogeographic, and evolutionary factors are connected by physiological and ecological interactions that promote biodiversity across the surface Limiting f
15	2012	WEB	...tionalgeographic.com	climate change, and meteor strikes have reduced the earth to small refuges that protect biodiversity before, and now we are just about to do the same. TI
16	1995	ACAD	Environment	, in many fields, science is inadequate. For example, efforts to preserve biodiversity are hampered because we do not know how many species there are or
17	1999	MAG	VegTimes	required to produce, distribute and apply synthetic fertilizers. Organic farming practices also promote biodiversity -the perpetuation of a broad variety of p
18	2007	MAG	Environmental	often have a top-down effect on habitats, helping to shape the landscape and preserve biodiversity . North America's big beasts weren't just an Ice Age anc
19	2011	ACAD	Bioscience	caused by the loss of symbionts further compromises the capacity of these reserves to conserve biodiversity over the long term. A substantial shift toward
20	2012	WEB	...nersdictionaries.com	protect biodiversity.? (Compare: We clearly need to do more to protect biodiversity .) It is unfortunate that the practice of keeping monkeys as pets still con

Fig. 12: Concordance lines for verb + biodiversity (five most frequent verbs)

The warm-up demonstrates that one-to-one translations using dictionaries do not necessarily yield the right results but that research on actual language use in context is needed. By introducing and discussing the concordance lines, you not only provide a first example of data-driven learning, but also give students a taste of what they will be working with during the upcoming lessons. However, you should point out that this list of examples is not representative and does not allow for an immediate conclusion as to which verb is most suitable for translating *Wir müssen die Artenvielfalt erhalten*. The final answer will be given in the warm-up activity of the next lesson.

Preparation

Before students can start working with the corpus, they need to register on the website and create a virtual corpus. For this purpose, hand out Worksheet A and guide the learners through the instructions (“Preparation of our corpus work”).

Ask a student to come to the front to create their account and the virtual corpus on the interactive whiteboard while giving them the necessary instructions. In the meantime, the other students follow those steps on their own computers.

Practice

As preparation for their task, show your students how to use the “List”-function by going through the steps on the whiteboard while they follow them on their computers. Just as in the preparation phase, use the word *sustainability* as an example to show how the “List”-function works.

Explain that the resulting concordance lines are the most fundamental function of corpus work, since each of these concordance lines displays a different, authentic, contextual use of the word *sustainability*.

After this introduction, students can continue with the practice tasks on the worksheet either individually or in pairs. Note that the worksheet includes step-by-step instructions as additional support.

The task requires students to implement what they have learnt about corpus work by searching for nouns that occur with *sustainable* and aims to expand their vocabulary to improve their language for this project. To this end, students will be asked to write their own sentences using the word combinations they found. Tell your students that they can use these or similar sentences for their project later on.

Use the remaining five to ten minutes of the lesson for the students to share their results (both the words and example sentences from the corpus, and their own sentences) and to talk about and answer issues and questions.

3.3 Lesson 2

Teacher preparation

Collecting vocabulary for the mind map

As an introduction to the lesson students create a mind map of their existing vocabulary knowledge which could be relevant for their chosen topics. The mind map should display the topics that the groups have chosen (Fig. 13).

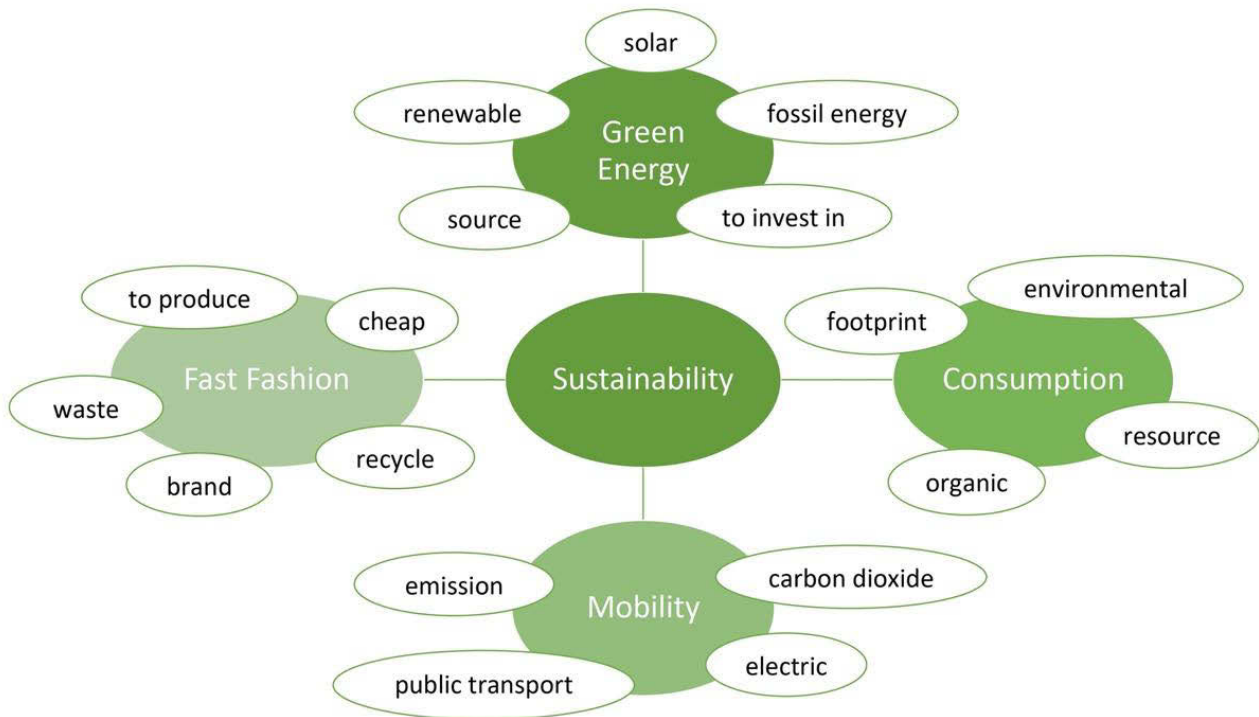


Fig. 13: Example mind map

You can prepare the main skeleton of the mind map (*sustainability* in the middle and the topics of the group projects) on the board or, if available, on an interactive whiteboard before the lesson.

Use the teaching material from previous lessons to note down vocabulary that you can expect students to come up with for each of their topics. Additionally, use the corpus to collect more key vocabulary and further ideas to help you guide the discussion in class and to make sure that the key terms for each topic are included in the mind map.

To browse for keywords in the virtual corpus, go to www.english-corpora.org and ensure that you are logged into your account. Then, go to the COCA again and select "Texts/Virtual." Click on "Edit corpora" to access your virtual corpus (Fig. 14).

Corpus of Contemporary American English

SEARCH VIRTUAL CORPORA CONTEXT OVERVIEW

[List](#) [Chart](#) [Word](#) [Browse](#) +

[POS] ?
[Find matching strings](#) [Reset](#)

Sections **Texts/Virtual** [Sort/Limit](#) [Options](#)

FIND TEXTS
 MY CORPORA

 sustainability

[Create corpus](#)
[Edit corpora](#)
[Find keywords](#)
[Refresh list](#)

(HIDE HELP) [Logout](#)

VIRTUAL CORPORA
 Create a "virtual corpus" -- essentially your own personalized corpus within COCA. You can create the corpus either by **keywords in the texts** (e.g. texts with the words *investments*, *basketball*, or *biology*), or **information about the texts** (e.g. date, title, or source), or a combination of keyword and text information.

You can then **edit** your virtual corpora, **search** within a particular virtual corpus, **compare** the frequency of a word, phrase or grammatical construction in your different virtual corpora, and also create "**keyword lists**" based on the texts in your virtual corpus.

Click on any of the links above for more information.

[Detailed overview \(NEW: Aug 2020\)](#)

Fig. 14: Edit corpora

You now see the page that contains information about your corpus. In the column "FIND KEYWORDS", click on "NOUNS" (Fig. 15) to obtain a list of nouns that are particularly frequent in your corpus as compared to a large reference corpus of general English (Fig. 16).

Corpus of Contemporary American English

SEARCH VIRTUAL CORPORA CONTEXT OVERVIEW

MY VIRTUAL CORPORA [Detailed overview \(NEW: Aug 2020\)](#)

HELP		↑	↓	LIST NAME ↑	# TEXTS ↓	# WORDS ↓	FIND KEYWORDS	<input checked="" type="radio"/> SPECIFIC	<input type="radio"/> FREQ	CREATED ↓
1				SUSTAINABILITY	746	3,662,544	NOUN VERB ADJ ADV N+N ADJ+N			175 d

Fig. 15: Finding keywords

Choose a different part of speech. It would be helpful for students to have different nouns, verbs, adjectives and adverbs at hand.

Here you have the option to modify your search (+ or - → „SPECIFIC“) to make it more or less specific to your corpus. If you make your search less specific, words will be sorted according to their absolute frequency, and you will obtain results such as *be* and *have* for verbs, and *thing* and *people* for nouns. By specifying your search, corpus-specific words like *emission* and *denier* will appear higher on the list, making it more relevant to our search.

Browse through the results to find those relevant for the topics your students have chosen.

HELP	WORD (CLICK FOR CONTEXT)	FREQ	# TEXTS	SPECIFIC	ENTIRE CORPUS	EXPECTED
1	ONE	690	203	FREQ 90 74 TEXTS	1,240	4.6
2	GOOD	434	121		895	3.3
3	SUSTAINABILITY	110				
4	EMISSION					
5	SUBSIDY					
6	MEAN					
7	SAVING					
8	RESOURCE					
9	FOSSIL					
10	COAL					
11	FARMER					
12	SCIENTIST					
13	FUEL					

Fig. 16: Keyword List

Do have in mind that your students will search for collocations during the lesson to expand their vocabulary in this way, meaning the words on the mind map should predominantly be words they have come across before.

When you find words in the keywords list, you can also click on the word to see the concordance lines for this word. This might be helpful to find out in what contexts a word is used, and to make sure that the word is relevant to your students' topics.

Tasks and instructions

Plan for the second lesson

Duration: 90 minutes

Lesson plan for lesson 2

Phase	Task	Implementation	Interaction form	Materials
(1) Warm-Up 5 min	Translating: <i>Wir müssen die Artenvielfalt erhalten.</i>	Presenting results of the "Collocates"-function for verb + <i>biodiversity</i> Discussion of results	Plenum	Electronic devices*
(2) Introduction I 10 min	Mind mapping	Finding vocabulary for topics on <i>sustainability</i>	Plenum	Electronic devices, textbook
(3) Introduction II 10 min	Vocabulary guide (1)	Deciding on keywords for blogs Filling in the "Draft structure" and "Keywords" column	Project groups	Electronic devices, Worksheet C
(4) Preparation 15-20 min	Getting to know the "Collocates"- function	Using the step-by-step guide on the worksheet to find noun collocates of <i>sustainable</i> in the corpus	Individual/ partner work	Electronic devices, Worksheet B
(5) Practice 35-45 min	Working on the project: Vocabulary guide (2)	Researching suitable vocabulary for blogs Completing the vocabulary guide by using the "Collocates"- function	Project groups	Electronic devices, Worksheet B, Worksheet C
(6) Conclusion and reflection 5-10 min	What did we learn? What can we do better?	– Discussion of ideas and questions; providing outlook of next lessons	Plenum	—

* = interactive whiteboard, students' computers/laptops

Warm-up

Begin again by presenting your sentence (*Wir müssen die Artenvielfalt erhalten.*) on the whiteboard and asking the learners: *What can you say about the translation of this sentence into English?* The following class-discussion partly recaps the content of the last lesson (problem of one-to-one translations, importance of corpus work and collocations). You can now tell them how you used the "Collocates"-function to come up with a list of verbs which are most frequently used in front of *biodiversity*. Show them the screenshot (Fig. 17) and discuss the results with your students.

Corpus of Contemporary American English

SEARCH FREQUENCY CONTEXT OVERVIEW

ON CLICK: CONTEXT TRANSLATE (??) GOOGLE IMAGE PRON/VIDEO BOOK (HELP)

WORD PROFILES: BIODIVERSITY

HELP		CONTEXT	FREQ
1	<input checked="" type="checkbox"/>	[CONSERVE]	83
2	<input checked="" type="checkbox"/>	[PROTECT]	74
3	<input checked="" type="checkbox"/>	[PRESERVE]	29
4	<input checked="" type="checkbox"/>	[REDUCE]	28
5	<input checked="" type="checkbox"/>	[PROMOTE]	25
6	<input type="checkbox"/>	[MAINTAIN]	23
7	<input type="checkbox"/>	[SEE]	17
8	<input type="checkbox"/>	[BE]	16
9	<input type="checkbox"/>	[THREATEN]	14
10	<input type="checkbox"/>	[AFFECT]	13
11	<input type="checkbox"/>	[ENHANCE]	10
12	<input type="checkbox"/>	[ALTER]	9
13	<input type="checkbox"/>	[SAVE]	8
14	<input type="checkbox"/>	[INCREASE]	7
15	<input type="checkbox"/>	[DESTROY]	7
16	<input type="checkbox"/>	[SUSTAIN]	7
17	<input type="checkbox"/>	[CONSIDER]	6
18	<input type="checkbox"/>	[INTEGRATE]	6
19	<input type="checkbox"/>	[INVESTIGATE]	6
20	<input type="checkbox"/>	[INFLUENCE]	6

Fig. 17: List of verb collocates of biodiversity

At least in American English, *preserve* is used much more often than *maintain*. As a matter of fact, it is the verb *conserve* that is most likely to occur in front of *biodiversity*, which is surprising given the fact that it was not mentioned at all in the dictionary entry on www.pons.de.

Use this example to briefly explain what collocations are: though *conserve*, *preserve* and *maintain* all translate as *erhalten* in German, they cannot be used interchangeably. The verb that is most frequently used with *biodiversity* appears to be *conserve*. That means that *biodiversity* collocates most strongly with *conserve*. This is what we call a collocation: words that are commonly used together.

Introduction (mind mapping)

Present the prepared mind map to the students, and ask them to think about vocabulary for each of the topics. Stress that **all students** should think about vocabulary for **all topics**. This way, every group can receive support and inspiration from their peers. Use the vocabulary collected in preparation for the lesson to guide the discussion so that your students collect a variety of words for different contexts, and also use it to add important words that your students do not mention. To provide more context, ask the groups to briefly state the arguments and ideas they want to address in their project to ensure that the collected vocabulary is relevant to the groups' project idea.

Introduction (vocabulary guide)

After the class discussion, ask your students to get together in their project groups, and hand out Worksheet C (the worksheet should either be printed in an A3 format or made available as a document on their computers). Explain that this worksheet is a vocabulary guide that the students will devise themselves to prepare for elaborating their products.

Ask the students to fill in the “Draft structure” column with a brief version of the draft structure they have developed in earlier lessons. Afterwards, they should use the mind map to decide on six keywords that are most central to their ideas and topics, and write them in the “Keywords” column (Fig. 18). Ask them to include nouns as well as verbs and adjectives.

Topic: Mobility Group Members: _____

Draft structure	Keywords	Collocations	Example Sentence 1	Example Sentence 2
1. traffic is bad for the environment	emission			
	public transport			
2. use public transport more often	to reduce			
	pollution			
3. try car sharing				
4. cycling				

Fig. 18: Example table for the introductory task (worksheet C)

Preparation

After the groups have completed the second column, students should turn to their computers or laptops. Hand out Worksheet B and go through it together. Make sure that all students have a basic understanding of what collocations are, then guide them through the step-by-step instructions while demonstrating the steps on www.english-corpora.org on screen, using *sustainability* as an example (Fig. 19).



Fig. 19: Search form for the collocation task

Help your students to understand how the results page (Fig. 20) works as it will look confusing at first: The words are grouped according to their part of speech (POS) and the colours refer to their frequency. The boxes to the left above the table display the word they searched for and its part of speech (e.g., noun). If they have searched for a word that could be more than one POS, they would find an option to switch the POS right next to it (e.g., *make* can be both a verb and a noun).

COLLOCATES		SUSTAINABILITY	NOUN	Advanced options				Collocates	Clusters	Topics	Dictionary	KWIC
+ NOUN	NEW WORD	?	+ ADJ	NEW WORD	?	+ VERB	NEW WORD	?	+ ADV	NEW WORD	?	
184	3.13	issue	472	6.70	environmental	113	4.72	achieve	15	2.56	increasingly	
138	2.69	program	214	6.14	corporate	92	4.79	ensure	10	5.99	eg	
105	3.44	energy	187	6.23	long-term	79	4.31	promote	8	4.49	ie	
95	3.18	development	162	3.29	social	56	2.66	focus	7	4.65	environmentally	
93	3.54	goal	144	3.89	economic	55	3.34	improve	4	2.68	namely	
92	3.14	effort	118	4.37	global	53	3.30	address	4	5.71	inextricably	
86	3.89	environment	107	5.37	urban	47	4.36	assess	3	2.50	nationwide	
80	3.01	practice	85	7.08	ecological	44	3.28	define	3	2.67	per	
79	3.57	resource	64	3.25	financial	41	2.91	relate	3	3.44	se	
77	2.97	science	57	3.69	cultural	38	4.45	enhance	2	2.55	internationally	
77	3.86	management	50	5.12	fiscal	34	4.94	integrate	2	2.67	systematically	
75	3.67	challenge	49	3.43	future	31	2.58	maintain	2	2.79	globally	
72	5.08	initiative	42	2.87	green	30	4.29	incorporate	2	2.97	vigorously	
68	2.71	project	42	5.51	sustainable	25	2.64	threaten	2	3.01	universally	
66	3.44	strategy	31	4.94	agricultural	25	3.40	implement	2	4.81	ecologically	

Fig. 20: Results page

Note that students should **not** click on the words but on the text icon next to the words to see the concordance lines (Fig. 21). Clicking on the word itself redirects to the collocates of that particular word. If you then go back to the previous page and request the concordance lines again, the interface will show you the concordance lines for the word you clicked on before rather than the concordance lines for *sustainability* (your search word). To fix this error, you need to launch your search again.

Tip: It might help to ask your students to ensure that two different words are highlighted in the concordance lines, one of which should be their search term. If that is not the case, the aforementioned error has been made and the students need to go back to the search form.

When going to the concordance lines, do not forget to randomise your search by clicking on a number after “Find sample”.

After going through the instructions, students form pairs and search for collocations of *sustainable* with the help of the step-by-step-guide to practise using the “Collocates”-function.

Tell your students to use the checklist on Worksheet B as a guideline to help them choose suitable collocations: Point out that students should look for new words rather than words that they have used already in the lesson before, i.e., focusing on verbs, adjectives and adverbs rather than nouns. They should also check that they understand how to use the word by checking the example sentences from the concordance lines. By writing a sentence for each collocation themselves, students can test whether they have really understood how to use the collocation in context.

Practice: Working on the project

Ask your students to get back together in their project groups and use the step-by-step guide, as well as the checklist on their worksheet (Worksheet B), to search for collocations for the keywords they have chosen (Worksheet C, Fig. 21). They can refer to the tasks on *sustainable* for support. Students then work on completing the table in their groups. Ask your students to underline the construction in the example sentences in which the collocation occurs.

Worksheet C: A Vocabulary Guide for Your Project

Sustainability – Act Responsible, Think Sustainable

Topic: Mobility Group Members: _____

Draft structure	Keywords	Collocations	Example Sentence 1	Example Sentence 2
1. traffic is bad for the environment	emission	control	...and <u>control wastes and emissions</u> more efficiently.	Republicans need to come up with a plan to <u>control carbon emissions</u> ?
		significantly	aimed to <u>significantly cut carbon emissions</u> .	the United States will not <u>significantly decrease carbon emissions</u>
2. use public transport more often	public transport	infrastructure	access to local destinations and <u>public transport infrastructure</u> .	
	to reduce			

Fig. 21: Worksheet C

After finishing the table, students can use the new words and example sentences to write their own sentences with the collocations or to start the first paragraph of their final project. Students need to apply the collocation in a construction that mirrors the one from the respective example sentence.

Finish the lesson with a class discussion on the results of the group work.

4 Options and further ideas

Topics

The topic chosen in this chapter is exemplary and could be narrowed down for more support, or replaced by another subject that is deemed relevant and suitable for the students, e.g.:

- Bullying: cyber bullying, peer pressure, ...
- Social networks: anonymity, abusive language, benefits, risks, ...
- Corona crisis: threats, vulnerabilities, measures, public health policies, economic impact, face masks, ...

Project outcomes

The proposed lessons can be used to prepare for various project tasks, e.g.:

- Creating a video or podcast
- Writing a blog entry
- Staging an interview
- Doing a role play / creating a sketch

Differentiation

The worksheets and the step-by-step guide will provide enough scaffolding for most students to be able to work on the tasks independently in their groups. However, for more support you could reduce the number of collocations and/or the number of keywords for individual groups or the whole class, or fill in parts of the vocabulary guide (Worksheet C), e.g., the “Collocations” column, and leave the rest of the table for the students to complete it, e.g., the “Example sentences” column.

Additional tasks

During the writing process, the “Collocates”-function can be continuously used to help students write their texts, e.g., by looking up collocations during the writing phase when they are unsure how to use words correctly, using the “Collocates”-function to improve their first drafts: Ask your students to read through

their own (or another group's) text and extract a certain number of phrases which they would like to improve. Then ask them to find collocations for these words to express their idea in a more idiomatic way.

5 Caveats and limitations

The lessons depend heavily on the availability of end devices for all students; it is assumed that each student has a computer, laptop or tablet they can use in class. If that is not the case, you can adjust the exercises by forming groups for the tasks.

Moreover, a secure and stable internet connection is necessary for these tasks, which the teacher should test in advance. Nonetheless, it is possible that the website gets overloaded if all students simultaneously search for the same things. This can be solved by forming pairs or groups that work on one computer at a time.

6 Conclusion

Although the outline of this project may appear labour-intensive at first, it sets the basis for further corpus-based learning and is therefore quite sustainable and time-saving in the long run. Once the teacher and students have acquired the necessary technical skills, they can be used again and expanded in future projects on any topic. Most importantly, students not only learn how to obtain linguistic support for their writing, but also develop a deeper understanding of how language works, i.e., in patterns rather than by the combination of individual words. Learning new words and phrases in context by actively and autonomously searching for collocations that suit their needs, students are likely to remember the new words better and to use them idiomatically. In this way, the designed lessons are in line with current evidence from second language acquisition research, which highlights the importance of rich exposure, noticing, learner relevance, and emotional as well as cognitive engagement.

7 Resources and references

Corpus of Contemporary American English (COCA). <https://www.english-corpora.org/coca/>.

Niedersächsisches Kultusministerium. 2015. Kerncurriculum für die integrierte Gesamtschule, Schulgänge 5-10: Englisch. <https://cuvo.nibis.de/cuvo.php?p=download&upload=66>.

Figures used on Worksheet A: <https://www.digitale-nachhaltigkeit.ch/icons-voraussetzungen/> and <https://www.english-corpora.org/images/words.jpg>

Figure used on Worksheet B: <https://pixnio.com/objects/toys/puzzle-pieces-symbol-connect-hand-jigsaw>

8 Appendix

Worksheet A

Name: _____

How to talk about *sustainability* using authentic language! (Worksheet A)



In this worksheet, you will learn how to enrich your blog by finding suitable vocabulary for the topic you want to write about. For this, you will work with a so-called "corpus" on the website www.english-corpora.org.

Preparation of our corpus work

Before we get to know the different functions of the website, we have to prepare our corpus work a little. So, let's start!

Step 1 - Creating an account

1. Open www.english-corpora.org in your web browser. On the top menu bar, go to "my account" and select "Register".
2. Fill in the blanks: Type in your school's email address, choose "**GERMANY**" as your country, and select "**Other**" under "Category". Tick the box next to "I agree to the Terms and Conditions for this website", and finally click on "submit".
3. You should then receive a confirmation email to verify your account. Go to your school's email account, open the new email, and click on the link to finalise your registration.

Step 2 - Creating a virtual corpus (or a specialised corpus)

1. Go back to the home page of www.english-corpora.org. You should now be logged into your new account.
2. You now see a list of corpora (= plural of *corpus*) which you can choose from. Select the Corpus of Contemporary American English (COCA).
3. Underneath the text box, you will find "Texts/Virtual". Click on this section, and then choose "Create Corpus".
4. Although the new appearing web page might look confusing at first, the next steps are just as easy as the others:
Under "BLOG" in the "Genre/domain" section, choose every genre **except** "Acad" and "Legal" (since we want to concentrate on the language of blogs but excluding academic and legal texts). Then type in **sustainability** in the search field for "Words in text", and click on "submit".
5. You can save your findings right above the list of blogs that is returned. We want to call our new corpus "sustainability", so this is what we should type in the respective field (-> "SAVE AS").

Well done, you have now completed all necessary steps to start working with the corpus! 😊

This worksheet was downloaded from: Stahnke, Rieke, Marlena Wegesin & Lara Brede (2021). "Act responsible, think sustainable – Expanding vocabulary to write about *sustainability*." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

Name: _____

Practice: Getting to know www.english-corpora.org and its functions



Concordance lines and the "List"-function

The first method to look for new words and phrases for your project is a very basic function of the website.

1. Since the COCA (= Corpus of Contemporary American English) is still selected, you simply click on "SEARCH", and select the "List"-function (next to "Chart", "Word" and "Browse").
2. In this example, we want to know more about the use and the context of the word **sustainable**, which you should type into the text field. Before we submit our search, we need to select our own specialised corpus. To do so, click on "Texts/Virtual" and choose the newly created "sustainability" corpus from the list. Afterwards, click on "Find matching strings".
3. You will now see a short table that tells you how often the word *sustainable* is found within our virtual (specialised) corpus.

Write down the frequency ("FREQ"): _____

4. When you click on "sustainable", you will obtain a long list of examples that contain the word *sustainable*. These examples are called **concordance lines**. They show how *sustainable* is used in **context**. Before we take a closer look at those examples, we need to randomise the results by clicking on a number next to "FIND SAMPLE" in the top left corner to get a better overview of the data.

Have a closer look at the concordance lines. Choose **2** examples in which *sustainable* is followed by a **noun**. Write down these two nouns: 🔍

1. _____
2. _____

This worksheet was downloaded from: Stahnke, Rieke, Marlena Wegesin & Lara Brede (2021). "Act responsible, think sustainable – Expanding vocabulary to write about *sustainability*." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

Name: _____

5. Now click on each of these two concordance lines and write down the full sentences in which the word *sustainable* occurs. ✍️

Look up any words you do not understand in a dictionary. 🔍

1. _____

2. _____

6. Choose **one** of the nouns and write down your own sentence that includes the pair (*sustainable* + your choice of noun). 📝

This worksheet was downloaded from: Stahnke, Rieke, Marlena Wegesin & Lara Brede (2021). "Act responsible, think sustainable – Expanding vocabulary to write about *sustainability*." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

Worksheet B

Name: _____

How can we use *sustainable* in a sentence? – Getting to know the “Collocates”-function (Worksheet B)



What are collocations?

Words often appear together with certain other words, and there are certain words that don't come together. For example, the word *sense* often appears with the word *make* – *make sense*, making it a collocation. The collocation function of www.english-corpora.org helps us to find words that often appear together with the particular word we are searching for. In this way, it helps us to use the right words together to make our language more authentic.

Step-by-Step-Guide:

1. Go to www.english-corpora.org and make sure you are logged in. Click on “Corpus of Contemporary American English (COCA)”.
2. In the blue box click on the left, click on “+” and then on “Collocates”.
3. In the first field, type in the word you want to find collocations for. Note that you can only type in one word (e.g. *public transport*: type in “transport”). Leave the second field blank.
4. In the grid below, click on “4” to the left and “4” to the right. All fields from 1 to 4 should be green now.
5. Click on “Texts/Virtual” (written in grey below the grid) and select your “sustainability” corpus.
6. Click on “Find collocates”.
7. A page with four tables will pop up: one each for nouns, verbs, adjectives and adverbs. The words you see here are words that often appear together with the word you searched for. Browse through the list and look for words that you think might be interesting for your project. **Do NOT click on the word!**
8. Once you find a word, click on the little text symbol in the left column of the table next to your word. Now you should see the concordance lines with two words highlighted in green in each line: the word you searched for and the word you chose from the list. Here, you can find examples of how to use the two words together.

Checklist: Finding the right collocation

- I did not know or use this collocation before (e.g., *emission* → not *carbon emission*; ✘)
 - The collocation is interesting and relevant to my project topic and I can use it in my writing.
 - I understand how this collocation can be used in a sentence (look at the example sentences to find out).
 - I feel confident that I can use this collocation myself.
- If you are in doubt, ask a classmate or your teacher!

This worksheet was downloaded from: Stahnke, Rieke, Marlena Wegesin & Lara Brede (2021). “Act responsible, think sustainable – Expanding vocabulary to write about *sustainability*.” In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefol.pressbooks.com>. CC-BY-NC 4.0.

Name: _____

How can we use *sustainable* in a sentence? – Getting to know the

“Collocates”-function – Tasks

1. Search for collocations of the word *sustainable*. Choose **two** collocations (that means two words from the lists) that you find interesting and that you would like to use in your project. Write down the English word and the translation. 🔍

1. _____

2. _____

2. For each collocation / word, write down two example sentences from the concordance lines (remember to click on the text symbol next to the word, not on the word itself!) that show you how you can use the word in a sentence. ✍️

1. _____

2. _____

3. Write down one sentence with each collocation yourself. 📝

This worksheet was downloaded from: Stahnke, Rieke, Marlena Wegesin & Lara Brede (2021). "Act responsible, think sustainable – Expanding vocabulary to write about sustainability." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefull.pressbooks.com>. CC-BY-NC 4.0.

Worksheet C

Worksheet C: A vocabulary guide for your project

Topic: _____ Group Members: _____

Draft structure	Keywords	Collocations	Example sentence 1	Example sentence 2
1.				
2.				
3.				
4.				

This worksheet was downloaded from: Stahnke, Rieke, Marlena Wegesin & Lara Brede (2021). "Act responsible, think sustainable – Expanding vocabulary to write about sustainability." In Le Foll, Elen (Ed.), *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://elenlefull.pressbooks.com>. CC-BY-NC 4.0.

13. Building a corpus to create materials for a debate on pollution

MARIE GRÖNINGER AND HA MI NGUYEN

1 Introduction and rationale

Enabling learners of English to debate current issues that dominate public discourse is a key communicative aim of teaching English. As a major vector of climate change, pollution makes for an ideal debating topic for the English language classroom. However, given that it is a specialised topic, it can be difficult for teachers and learners to access the appropriate vocabulary and idiomatic phrases to do so successfully. The corpus built in this step-by-step guide to conducting corpus linguistics research helps teachers to compile lists of authentic articles and useful idiomatic phrases on the topic of *pollution*. The chapter will also introduce examples of how to implement the corpus findings in teaching materials. According to the English curriculum of Lower Saxony (Germany), dealing with topics from different school subjects in bilingual teaching is a highly motivating and effective way of acquiring a second language (Niedersächsisches Kultusministerium 2015: 31). Climate change is one of the core topics featured in the geography and chemistry curriculum (Niedersächsisches Kultusministerium 2015: 14) for grades 9 and 10 at Gymnasium (advanced secondary school in Germany), whilst also touching on contents from biology and politics.

Interdisciplinary climate change education “helps people understand and address the impact of global warming, increases ‘climate literacy’¹ among young people, encourages changes in their attitudes and behaviour, and helps them adapt to climate change related trends” (Unesco, n.d.). Pollution is a serious global issue that is addressed in political discussions. Therefore, the topic is very prevalent in students’ daily private lives. Many young people attend the *Fridays for Future* demonstrations, thus potentially increasing the students’ interest in the topic. Enabling students to communicate central issues of *pollution* in English is thus a meaningful preparation for encounters outside the classroom.

Outline and objectives

In this lesson, students will be introduced to general vocabulary around the topic of *pollution*. They will describe pictures, play a vocabulary game, and prepare and hold a debate on the topic of pollution. For the debate, the students will be divided into five groups. Four groups will be assigned a role and the members of the fifth group will moderate the debate. The four groups will read different articles which are related to their role, and work out arguments regarding *pollution* in order to debate with the opposing group.

Lesson time: ca. 90 min.

Teacher preparation time: ca. 60 min.

School Type: Gymnasium (upper secondary school)

Level: 10th Grade

Learning Objectives:

- Students can discuss ideas, opinions and knowledge about pollution as part of a roleplay.
- Students can formulate their own opinions on the topic.
- Students can make use of debating strategies and useful phrases for debating.
- Students can connect, describe and explain the causes and effects of pollution as a subtopic of climate change.
- Students can use key vocabulary pertaining to the semantic field of pollution.

2 Corpus, tools and methods

The News on the Web (NOW) corpus on english-corpora.org includes a range of broadsheet newspapers such as *The Guardian*, as well as tabloid press such as the *Daily Mail*. The corpus thus reflects the language input of the general public because there is a variety of written articles on the topic composed for different audiences. It is a rich source of frequent vocabulary useful for talking about scientific and political issues, as well as idiomatic phrases demonstrating how to use the new vocabulary in context. This kind of authentic language is very important for designing materials for the English as a Foreign Language classroom.

3 Step-by-step guide

Teacher preparation

Corpus compilation

Step 1: Registering on english-corpora.org

To be able to access the NOW corpus, register on english-corpora.org and select the corpus by clicking on “News on the Web (NOW)” in the table.

Corpus (online access)	Download	# words	Dialect	Time period	Genre(s)
iWeb: The Intelligent Web-based Corpus		14 billion	6 countries	2017	Web
News on the Web (NOW)		10.4 billion+	20 countries	2010-yesterday	Web: News
NEW The Coronavirus Corpus		429 million+	20 countries	Jan 2020-yesterday	Web: News
Global Web-Based English (GloWbE)		1.9 billion	20 countries	2012-13	Web (incl blogs)
Wikipedia Corpus		1.9 billion	(Various)	2014	Wikipedia
Corpus of Contemporary American English (COCA)		1.0 billion	American	1990-2019	Balanced
Corpus of Historical American English (COHA)		400 million	American	1810-2009	Balanced
The TV Corpus		325 million	6 countries	1950-2018	TV shows

Fig. 1: Accessing the News on the Web corpus through english-corpora.org

Step 2: Creating a virtual corpus

The “FIND TEXTS” function allows you to create a virtual corpus based on keywords. To use this function, type in your keyword in the query box; in this case it is *pollution*. Second, select *Texts/Virtual* tab and then click on “FIND TEXTS” (see Fig. 2).

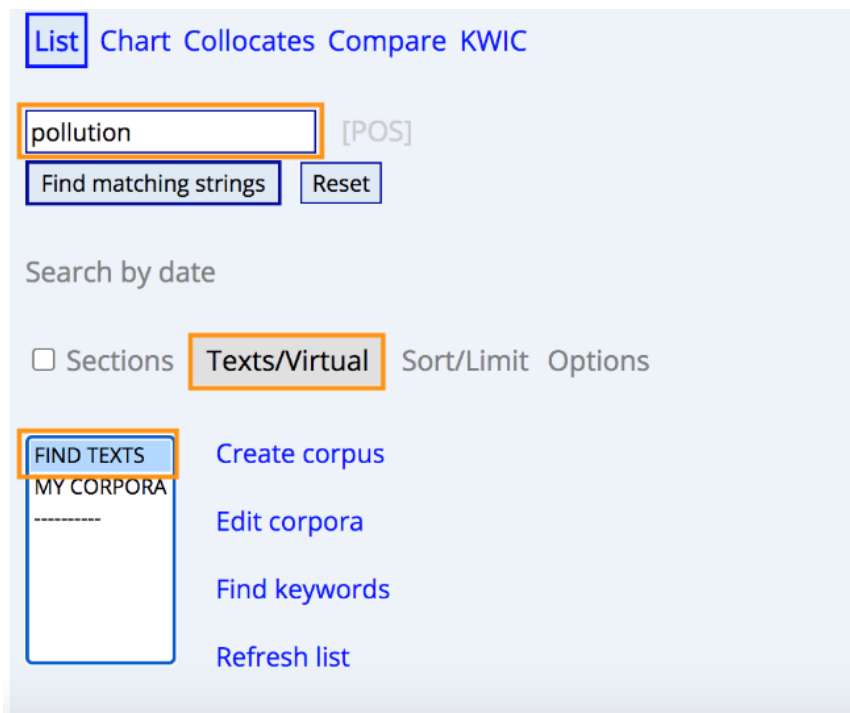


Fig. 2: Creating a virtual corpus with the keyword *pollution*

Third, select *Options* and type in your desired number of results in the “#HITS” box. To ensure a greater variety of texts, 1,000 texts could be retrieved. To see the results, select “Find matching strings” once you have adjusted the settings as described above (see Fig. 3).

List Chart Collocates Compare KWIC

[POS]

Search by date

Sections Texts/Virtual Sort/Limit

HITS

KWIC

GROUP BY

DISPLAY

SAVE LISTS

Fig. 3: Adjusting the settings of the virtual corpus

Next, you will see a list of texts that form your personal virtual corpus. You can click on the “SAVE LIST” option at the top left corner of the list to save your virtual corpus and to access it at any time in the future.

Designing corpus-based teaching materials

Task 1

To find certain keywords in your corpus, you can select e.g. “NOUN” to find specific nouns in your corpus. This option appears in your “virtual corpora” list next to the name of your saved corpus (see Fig. 4). If you click on “NOUN”, a list of nouns and their frequencies in your corpus will appear. If you then click on a word from this list, e.g. *emission* (see Fig. 5), a list is displayed showing in which context the word *emission* is used in the newspaper articles of your virtual corpus, alongside the date of the article, the website it was published on, and further information. You can also click on the number in the left column to see the expanded context in which the word is being used (see Fig. 6).

MY VIRTUAL CORPORA [Detailed overview \(NEW: Aug 2020\)](#)

HELP	↓	↓	LIST NAME ↓	# TEXTS ↓	# WORDS ↓	FIND KEYWORDS ○ SPECIFIC ● FREQ	CREATED ↓
1			POLLUTION	1000	2,498,244	NOUN VERB ADJ ADV N+N ADJ+N	187 d

Fig. 4: “Find keywords” option

HELP	WORD (CLICK FOR CONTEXT)	FREQ	# TEXTS	SPECIFIC		ENTIRE CORPUS	EXPECTED
				FREQ	TEXTS		
1	POLLUTION	17726	915	1.3		65,208	13,805.6
2	AIR	13229	838	0.1		491,484	104,054.8
3	YEAR	7050	801	0.0		6,850,677	1,450,395.1
4	GOVERNMENT	4677	546	0.0		2,376,887	503,224.0
5	LEVEL	4574	790	0.0		872,811	184,787.7
6	PEOPLE	4418	721	0.0		4,745,533	1,004,703.3
7	COUNTRY	4111	585	0.0		2,128,105	450,553.0
8	CITY	3830	619	0.0		1,829,468	387,326.9
9	TIME	3652	694	0.0		4,951,917	1,048,398.0
10	HEALTH	3577	728	0.0		1,028,503	217,750.1
11	EMISSION	3318	514	0.2		104,135	22,047.0

Fig. 5: List of nouns

CLICK FOR MORE CONTEXT		<input type="checkbox"/> [?]	SAVE LIST	CHOOSE LIST	CREATE NEW LIST	<input type="checkbox"/> [?]	SHOW DUPLICATES
1	20-07-13 KE	citizentv.co.ke x	A	B	C	cleaning system was within legal limits. # But it turned out that not all emissions were being captured in their observations and a huge plume of dark smoke	
2	20-07-13 KE	citizentv.co.ke x	A	B	C	Endmor Steel Millers was polluting their neighborhood. # Since the factory wasn't releasing emissions through its chimney, he captured most of the smoke e	
3	20-07-13 KE	citizentv.co.ke x	A	B	C	ruled that Endmor Steel Millers desist from any further pollution. The tribunal demanded an emissions report from Endmor. # The hearing of the main appe	
4	20-07-13 KE	citizentv.co.ke x	A	B	C	to the day then the temperatures are high... so at night it means the emissions will just fall in the almost immediate neighborhood because its cooler, " Ms. G	
5	20-07-13 KE	citizentv.co.ke x	A	B	C	yet to give a stack emission report, one measuring the contents of the actual emissions leaving the factory. # Were it not for the postponement of the March	
6	20-07-13 KE	citizentv.co.ke x	A	B	C	# As the sun went further to the West, things started to change: emissions from the factory started to be visible. # By the time we were doing	
7	20-07-13 KE	citizentv.co.ke x	A	B	C	neighborhood usually known as ambient air and that it was yet to give a stack emission report, one measuring the contents of the actual emissions leaving tr	
8	20-07-13 KE	citizentv.co.ke x	A	B	C	assessing its capacity and knowledge of the national environmental laws and regulations which provide the emission limits for these parameters in question	

Fig. 6: List of results for the noun emission

For the warm-up activity, which in this lesson will be the vocabulary game *Last Man Standing*, the teacher needs to prepare a list of vocabulary for the students to learn before this lesson.

This vocabulary list contains important vocabulary on the topic of *pollution*. The list will include the English word with three typical collocations and an example sentence in which the word is used.

To create this list, use the "FIND KEYWORDS" function and select as many nouns, verbs, adjectives etc. as needed for the vocabulary list (see Fig. 4).

There is going to be a huge collection of words to select from. It is therefore important to choose them carefully with your students in mind. First of all, you can find the most frequent nouns, verbs, adjectives etc. in your virtual corpus on *pollution* (see Fig. 4). Of course, not every word is going to be relevant for your students. It is helpful to view the text(s) that the students will read in the second task as a comparison in order to see which of the listed words may be necessary to understand them. Furthermore, Francine Johnston's article *How to Select Words for Vocabulary Instruction* provides handy criteria for the selection. Some of the criteria Johnston mentioned are "importance" and "transferability" of words (Johnston, n.d). Teachers should ask themselves what important words their students need to learn in order to understand the selection of texts and to successfully hold a debate on their own.

The next step after having chosen the appropriate vocabulary is to find two to three collocations per word. To start off, e.g. type in *emission* in your virtual *pollution* corpus (from 4.1.1) and choose the function "Collocates" (see Fig. 7).

The screenshot shows the iWeb interface for finding collocations. The search box contains "emission" and the second box contains "VERB". The collocate box contains "verb.ALL". A navigation bar at the top has tabs for SEARCH, KEYWORDS, CONTEXT, and TEXTS. The "Collocates" tab is active. On the right, there is a "VIRTUAL CORPORA" section with instructions on how to create and use virtual corpora.

Fig. 7: Finding collocations

Under *Options*, is also helpful to select "GROUP BY: LEMMAS" instead of the default "WORDS" (see Fig. 8) because the different inflections of verbs and nouns are then grouped together.

The screenshot shows the iWeb interface for finding collocations with options. The search box contains "emission" and the second box contains "VERB". The collocate box contains "verb.ALL". The "Options" tab is active. The "GROUP BY" dropdown is set to "LEMMAS". On the right, there is an "OTHER OPTIONS" section with explanations for # HITS, # KWIC, GROUP BY, DISPLAY, and SAVE LISTS.

Fig. 8: Group by lemmas

There are several options to look for useful collocations of the word. For example, you can type in an asterisk (*) in the second search box to look for all word classes. Another option is to specifically search for a word class such as verbs in order to find out which verbs are frequently associated with the word in texts about pollution (see Fig. 8). It may be necessary to vary the number of positions before or after the word depending on what word class you are looking for. Verbs, for example, are often several word positions before a noun.

You can change the collocational range you are searching for by adjusting the numbers between +4 and 4+. The range queried is highlighted in green (see Fig. 7).

In this case, the verb *meet* was the most frequent verb collocate of *emission* within our virtual *pollution* corpus (see Fig. 9).

ON CLICK: [CONTEXT](#) [TRANSLATE \(??\)](#) [GOOGLE](#) [IMAGE](#) [PRON/VIDEO](#) [BOOK](#) (HELP)

HELP	<input type="checkbox"/> ?	CONTEXT	FREQ
1	<input type="checkbox"/>	[MEET]	26
2	<input type="checkbox"/>	[CONTROL]	16
3	<input type="checkbox"/>	[SET]	15
4	<input type="checkbox"/>	[INSTALL]	14

Fig. 9: Selecting collocations

However, a brief look at the concordance lines of *meet* reveals that *meet + emission(s)* never occurs on its own (see Fig. 10). Instead, the verb *meet* collocates with *emission norms/standards/targets* and other semantically related nouns.

CLICK FOR MORE CONTEXT [?] [SAVE LIST](#) CHOOSE LIST CREATE NEW LIST [?] SHOW DUPLICATES

1	20-07-11 IN	downtoearth.org.in	A B C	. In iron ore preparation stage called sintering many plants were found not meeting stack emission norms. # In the subsequent stage of liquid iron making
2	20-05-11 ZA	biznews.com	A B C	install flue gas desulphurisation (FGD) plant at Medupi in order to meet SO2 emission requirements. It is estimated that FGD plant for Medupi will cost som
3	20-03-12 US	Federal Register	A B C	or # (3) Before September 9, 2020, fails to meet any emission limitation (including any operating limit) or work practice standard in this subpart during
4	20-03-12 US	Federal Register	A B C	the facility could use the performance test to demonstrate that it can still meet the emission limit beyond the maximum pressure drop recommended by t
5	20-01-19 IN	indiaspend.com	A B C	well within the standard of 300 mg/Nm3 and new technologies would help it meet the emission standards, the Business Standard reported in August 2019.
7	19-06-28 GH	Peace FM Online (1)	A B C	Driver's Vehicle license Authority could help remove vehicles, which do not meet emission and efficiency standards from Ghana's roads and an efficient pul
8	19-04-13 AU	Brisbane Times	A B C	by encouraging the take-up of replacement of all heaters with new ones that meet stricter emission and efficiency standards. Mortality risks # Ivan Hanigar
9	18-10-23 IN	Moneycontrol.com	A B C	allowed the sale of fireworks, but on the condition that such firecrackers meet low emission and decibel standards. The Hindustan Times reported that the
10	18-05-24 HK	Vantage Asia	A B C	as residential and commercial areas, to relocate their facilities unless they can meet strict emission limits. In principle, the plan requires all medium and sm
11	18-05-24 HK	Vantage Asia	A B C	the level of industrial parks, if the park as a whole does not meet emission targets, the government may order all park facilities to lower emissions or temp
12	17-11-16 IN	Scroll.in	A B C	low. In the current system, a computerised test decides if a vehicle meets emission standards and issues Pollution Under Control certificates that have to b

Fig. 10: Concordance lines for *meet*

When creating your own collocation list, please keep in mind that a certain inflection of a word may be more idiomatic than others. For example, while looking for *verb + emission* collocations, it becomes clear that the plural of *emission*, i.e. *emissions*, is used much more frequently. The same is true for *adjective* or *noun + emission*. The singular form of *emission*, on the other hand, is mostly found in *emission + noun* constructions, i. e. *emission standards*. As a tip, an easy way to check this is to search for different inflections in a corpus while using the setting "GROUP BY: WORDS". The frequency will help you to decide which inflections should be included in the collocations you decide to include in your teaching materials.

In addition, learners are known to find example sentences most useful. Make sure that they are from your corpus to ensure authenticity. An example of what a finished vocabulary list could look like can be found in the appendix.

Task 2: Instructions for one example text

The second task prepares the students for their roles in the debate. To do this, the teacher needs to prepare a variety of texts for each group in advance. The texts should inform the students about pollution and ideally contain useful information and arguments which can then be used in the debate. In the following, two ways of finding appropriate texts are outlined.

Option 1: Searching for a text in your own virtual corpus

For this lesson, this option was chosen to find the ideal articles for the reading task. The virtual corpus includes many great articles that suit the purpose of this lesson. To find a text in your individual corpus, select your virtual corpus. In this case, it is the *pollution* corpus that was created in 4.1.1. (see Fig. 11). Within this corpus, you have a variety of texts that you can choose from. Next, you can click through the websites to find the perfect text(s) for the lesson (see Fig. 12).

MY VIRTUAL CORPORA

HELP	↑	↓	LIST NAME ↓	# TEXTS ↓	# WORDS ↓	FIND KEYWORDS <input checked="" type="radio"/> SPECIFIC <input type="radio"/> FREQ	CREATED ↓
1			POLLUTION	1000	2,498,244	NOUN VERB ADJ ADV N+N ADJ+N	1 h

Fig. 11: List of virtual corpora

POLLUTION (RENAME) DELETE ADD TO MOVE TO --SELECT-- (SEE ALL VIRTUAL CORPORA) SHARE LIST ?					
HELP	<input type="checkbox"/> 100	DATE	COUNTRY	WEBSITE	TITLE
1	<input type="checkbox"/>	20-01-05		Zakir Naik	High Level of Air Pollution Associated with Poor B...
2	<input type="checkbox"/>	20-01-15		GreenBiz	Climate change has led to more temperature inversi...
3	<input type="checkbox"/>	20-03-10		Acsh.org	Air Pollution and Mortality - Measuring the Long a...
4	<input type="checkbox"/>	20-04-14		Air Quality News	Ricardo: an analysis of Covid-19 lockdown on UK lo...
5	<input type="checkbox"/>	20-05-10		Red Green Blue	Pollution kills 9 million people a year. How is th...
6	<input type="checkbox"/>	20-06-11		Climate Central	Tear Gas, Pollution, Wildfire Smoke: A Triple Thre...

Fig. 12: Selection of articles within virtual corpus on pollution

Option 2: Searching for a text in the NOW corpus

This second option enables you to search through the entire NOW corpus by simply typing in a keyword. In this example, the term *air pollution* is used. In addition, it is possible to specify the dates of publication for the texts in your search (see Fig. 13). The more recent the texts are, the more likely it is that the full text will still be available on the publisher's website.



[List](#) [Chart](#) [Collocates](#) [Compare](#) [KWIC](#)

[POS] ?

[-] START to END

Sections Texts/Virtual Sort/Limit Options

MY CORPORA

-
- Pollution

(HIDE HELP) 1000/1000

VIRTUAL CORPORA

Create a "virtual corpus" -- essentially your own personalized corpus within NOW. You can create the corpus either by **keywords in the texts** (e.g. texts with the words *investments*, *basketball*, or *biology*), or **information about the texts** (e.g. date, title, or source), or a combination of keyword and text information.

You can then **edit** your virtual corpora, **search** within a particular virtual corpus, **compare** the frequency of a word, phrase or grammatical construction in your different virtual corpora, and also create **"keyword lists"** based on the texts in your virtual corpus.

Click on any of the links above for more information.

[Detailed overview \(NEW: Aug 2020\)](#)

Fig. 13: Search settings in the NOW corpus

Next, you will see a list of websites that have published articles featuring the word or topic that you searched for. Once you have found an interesting website, you can click on the number on the left column to obtain further information about the context in which your query word is used. In this case, no. 58 was chosen, which is an article from *The Guardian* (see Fig. 14).

SEARCH	FIND TEXTS	CONTEXT	ACCOUNT
44	<input checked="" type="checkbox"/>	IN 19-09-11: INDIASPEND.COM	3801 26 6,840.3
45	<input checked="" type="checkbox"/>	US 20-01-05: ZAKIR NAIK	1935 26 13,436.7
46	<input checked="" type="checkbox"/>	US 19-07-23: PLOS	5680 26 4,577.5
47	<input checked="" type="checkbox"/>	IE 20-09-12: INTERESTINGENGINEERI...	2489 26 10,446.0
48	<input checked="" type="checkbox"/>	KE 20-08-06: CAPITALFM.CO.KE	929 26 27,987.1
49	<input checked="" type="checkbox"/>	PK 18-11-02: THE NEWS INTERNATIONAL	1032 26 25,193.8
50	<input checked="" type="checkbox"/>	ZA 20-11-05: DAILYMAVERICK.CO.ZA	1726 26 15,063.7
51	<input checked="" type="checkbox"/>	PK 18-11-01: THE NEWS INTERNATIONAL	1032 26 25,193.8
52	<input checked="" type="checkbox"/>	IN 18-12-06: THE INDIAN EXPRESS	848 26 30,660.4
53	<input checked="" type="checkbox"/>	IN 19-05-17: INDIASPEND	2213 26 11,748.8
54	<input checked="" type="checkbox"/>	BD 20-11-03: THEINDEPENDENTBD.COM	2044 25 12,230.9
55	<input checked="" type="checkbox"/>	PH 16-11-20: BUSINESS MIRROR	1355 25 18,450.2
56	<input checked="" type="checkbox"/>	US 19-11-06: FIRSTPOST.COM	4864 25 5,139.8
57	<input checked="" type="checkbox"/>	IE 17-10-19: IRISH TIMES	913 25 27,382.3
58	<input checked="" type="checkbox"/>	US 19-05-17: THE GUARDIAN	1374 25 18,195.1
59	<input checked="" type="checkbox"/>	UK 19-11-24: SUNDAVORSEVEDUK	1326 25 18,712.6

Fig. 14: Selection of articles within the NOW corpus

CLICK FOR MORE CONTEXT	SAVE LIST	CHOOSE LIST	CREATE NEW LIST	SHOW DUPLICATES
1	19-05-17	US	The Guardian	A B C

New analysis indicates 8.8m early deaths each year -- double earlier estimates -- making **air pollution** a bigger killer than tobacco smoking. # But the impact of different p

Fig. 15: Selected article (here: from the newspaper *The Guardian*)

You can now click on the number in the left column to see the expanded context in which the word is being used (see Fig. 15).

You will also find a link to the article (see Fig. 16).

Source information:

Source	https://www.theguardian.com/environment/ng-interactive/2019/may/17/air-pollution-may-be-damaging-every-organ-and-cell-in-the-body-finds-global-review
Country/date	US (19-05-17)
Title	Revealed: air pollution may be damaging 'every organ in the body'

Expanded context:

and from liver problems and bladder cancer to brittle bones and damaged skin. Fertility, fetuses and children are also affected by toxic air, the review found. # The systemic damage is the result of pollutants causing inflammation that then floods through the body and ultrafine particles being carried around the body by the bloodstream. # **Air pollution** is a " public health emergency ", according to the World Health Organization, with more than 90% of the global population enduring toxic outdoor air. New analysis indicates 8.8m early deaths each year -- double earlier estimates -- making **air pollution** a bigger killer than tobacco smoking. # But the impact of different pollutants on many ailments remains to be established, suggesting well-known heart and lung damage is only " the tip of the iceberg ". # " **Air pollution** can harm acutely, as well as chronically, potentially affecting every organ in the body, " conclude the scientists from the Forum of International Respiratory Societies in the two review papers, published in the journal *Chest*. " Ultrafine particles pass through the lungs, are readily picked up by

Fig. 16: Source information and expanded context

Four example texts we suggest could be used for this task are linked in the appendix, though your own selection of texts will, of course, depend on your students' proficiency levels, needs and interests.

Task 3

The final task of the session is the debate. For this task, it is necessary to prepare a list of useful phrases for engaging in a debate. Teachers should provide their students with the list so that they can study it in advance. It will guide the learners through the debate and support them if needed. You may adjust your list to your students' needs by, for example, removing the translation into students' first language and leaving an empty column instead. This leaves space for your students to fill in their own translation, or other notes that they may find useful.

At this point, the students are already prepared and have acquired information and arguments for their specific roles.

Lesson plan

Table 1: Lesson Plan

Phase	Task	Implementation	Interaction Form	Material/Medium
Introduction 10 min	Picture impulse on the topic <i>pollution</i>	Description of and comments on the picture	Plenum	Smart board, Picture (see appendix)
Task 1 5 min	Vocabulary game: <i>Last Man Standing</i>	Teacher names the theme <i>pollution</i> , students name vocabulary	Plenum (standing)	Ball
Task 2.1 30 min	Preparing roles for debate	All groups read through their texts and develop arguments with the information provided	Group work	Worksheets, extra sheet for notes
Task 2.2 10 min	Exchanging arguments regarding the motion	Collaboration with the other pro/con role	Group work	Notes, arguments from 2.1
Task 3 20 min	Debate in small groups Motion: <i>Preventing Pollution is not a Matter of Money</i>	Students discuss in debate groups (one student per role) under the guidance of a moderator	Jigsaw group work	Notes, arguments from 2.1
Conclusion 10 min	Reflection on debates	Students comment on the strategies of the debaters and state their own opinions on the topic.	Plenum	
Application/ Homework 5 min	Analysing a cartoon	Finish analysis as written homework.	Plenum	Cartoon (see appendix)

Instructions for tasks

Task 1: Vocabulary Game – Last Man Standing (warm-up activity)

The teacher names the topic, *pollution*, for which the students have already learnt the vocabulary for homework ahead of the lesson. At the beginning, all of the students have to stand up and the teacher throws a ball to a student who has to come up with a word that fits that topic. That student throws the ball to the next student who again has to come up with another word. Students who catch the ball and cannot come up with a word have to sit down. This game continues until there is only one student left standing. The last standing student is the winner of the game.

Task 2.1: Reading

The teacher presents the motion of the debate “Preventing Pollution Is Not a Matter of Money”. In the first stage of the *Jigsaw* method, the students are divided into five expert groups. Each group will be assigned one of the roles.

First of all, informative texts are handed out to the first four groups with potential arguments supporting the position they will have to defend. These texts include reading tasks to help them filter and develop arguments for the debate. Moreover, they will have to think about strategies to communicate these arguments in the debate. The fifth group is the group of moderators. It is best suited for high ability students. These students have access to all four texts in order to prepare for their role. They will be in charge of opening the debate, facilitating transitions between arguments, encouraging everyone to participate actively and asking questions whenever needed.

Roles in the debate: “Preventing Pollution Is Not a Matter of Money”

PRO GROUP

- Representative of an environmental organization
- Student who supports Fridays for Future

CON GROUP

- CEO of a successful car manufacturer
- Parent of four who cannot afford an environmentally-friendly lifestyle

MODERATOR GROUP

Task 2.2: Exchanging arguments

Students who have been assigned the same role form groups and review their arguments. This step encourages the students to think about possible ways to support each other’s arguments, thus strengthening their side of the motion. The students should also think about possible arguments from the opposing side and come up with potential rebuttals.

Task 3: Debate

For the debate, new discussion groups are formed; those discussion groups include at least one representative of each expert group. This means that each discussion group has a representative of each role plus a moderator who is going to be in charge of structuring as well as keeping the debate going. The purpose of this format is to encourage every student to participate actively.

4 Options and further ideas

Option:

The teacher could create a pro and con group for low ability learners who may obtain texts that are easier to understand and/or include additional information such as graphs, images or an additional vocabulary list.

Further idea:

Fish Bowl debate for the follow-up lesson:

Fishbowl is a strategy for organizing medium- to large-group discussions. Students are separated into an inner and outer circle. In the inner circle, or fishbowl, students have a discussion; students in the outer circle listen to the discussion and take notes (tolerance.org, n.d.).

A *Fish Bowl* debate gives students the opportunity to present their findings in front of the whole class while the observers note down the structure of the debate, as well as the vocabulary employed. That way, the teacher is able to retrace how well the *Jigsaw* classroom worked out and the students can use their results in a bigger context while revising the acquired vocabulary and debating strategies once more.

5 Caveats and limitations

In order to create such a lesson, it is important to become familiar with english-corpora.org and not lose your patience at the first mishap! Error messages, resulting e.g. from forgetting to put a space or ticking the wrong box, are certainly frustrating. Nonetheless, the results obtained from working with a corpus are very reliable and enable you to find vocabulary that goes over and beyond what textbooks include, thus expanding the vocabulary taught in a more authentic way. Textbooks and other lists might include vocabulary around specific topics such as pollution, but the corpus allows you to look for frequent idiomatic phrases and collocations and to explore the wider context in which they are used.

Another difficulty might occur for non-paying users of english-corpora.org, who regularly receive a notification to upgrade their account. This can be ignored. After a few seconds of waiting, the search can be continued. Additionally, when creating the virtual corpus, it is important to choose a keyword that is often used in the desired context. It is worth comparing results from different keywords because the use of a good keyword is crucial to create a reliable virtual corpus. Depending on the learning group, the complexity of the tasks and the debate can vary. At first, students may be overwhelmed by the debating task, or else they might find it highly motivating and will be very spontaneous in their contributions. It is important to consider the specific learning group, that the necessary vocabulary has been introduced beforehand, and that the students are familiar with the concept of such a debate.

6 Conclusion

Even though preparing the vocabulary list may take some time, in our opinion, the result is very rewarding because you end up with a list of frequent, authentic examples that can always be reused or extended in future lessons. Moreover, this method can be applied to a whole range of other topics. Once the teacher has worked with this method once, they will always be able to go back and repeat the steps, applying the method to a different topic. Teacher preparation time will considerably decrease once they have gotten used to the method.

7 Resources

News On The Web Corpus (NOW). <https://www.english-corpora.org/now/>.

8 References

Climate.gov. 2009. What is Climate Literacy? <https://www.climate.gov/teaching/essential-principles-climate-literacy/what-climate-science-literacy> (09 January, 2021).

Englisch Hilfen. n.d. Diskussionen auf Englisch – Wichtige Wendungen. <https://www.englisch-hilfen.de/words/diskussionen.htm> (31 July, 2020).

Johnston, Francine. n.d. How to Select Words for Vocabulary Instruction. https://assets.pearsonschool.com/asset_mgr/current/201640/Francine_ArticleForResearch.pdf (31 July, 2020).

Lehrerfortbildung-bw.de. n.d. Material 12: Debate vocabulary. https://lehrerfortbildung-bw.de/u_sprachlit/englisch/gym/bp2004/fb1/binnendiff/2_ue_mat/mat14/ (04 August, 2020).

Niedersächsisches Kultusministerium. 2019. Kerncurriculum für das Gymnasium, Schuljahrgänge 5-10: Chemie. https://cuvo.nibis.de/cuvo.php?skey_lev0_0=Schulbereich&svalue_lev0_0=Sek+I&skey_lev0_1=Fach&svalue_lev0_1=Chemie&skey_lev0_2=Dokumentenart&svalue_lev0_2=Kerncurriculum&docid=1061&p=detail_view (15 July, 2020).

Niedersächsisches Kultusministerium. 2015. Kerncurriculum für das Gymnasium, Schuljahrgänge 5-10: Erdkunde. https://cuvo.nibis.de/cuvo.php?skey_lev0_0=Schulbereich&svalue_lev0_0=Sek+I&skey_lev0_1=Fach&svalue_lev0_1=Erdkunde&skey_lev0_2=Dokumentenart&svalue_lev0_2=Kerncurriculum&docid=1112&p=detail_view (15 July, 2020).

Niedersächsisches Kultusministerium. 2015. Kerncurriculum für das Gymnasium, Schuljahrgänge 5-10: Englisch. https://cuvo.nibis.de/cuvo.php?skey_lev0_0=Schulbereich&svalue_lev0_0=Sek+I&skey_lev0_1=Fach&svalue_lev0_1=Englisch&skey_lev0_2=Dokumentenart&svalue_lev0_2=Kerncurriculum&docid=1197&p=detail_view (15 July, 2020).

Tolerance.org Website. n.d. Teaching Strategy – Fish Bowl. <https://www.tolerance.org/classroom-resources/teaching-strategies/community-inquiry/fishbowl#:~:text=Fishbowl%20is%20a%20strategy%20for,the%20discussion%20and%20take%20notes> (12 August, 2020).

Unesco Website. n.d. Climate Change Education and Awareness. <https://en.unesco.org/themes/addressing-climate-change/climate-change-education-and->

9 Appendix

Pictures

Picture impulse

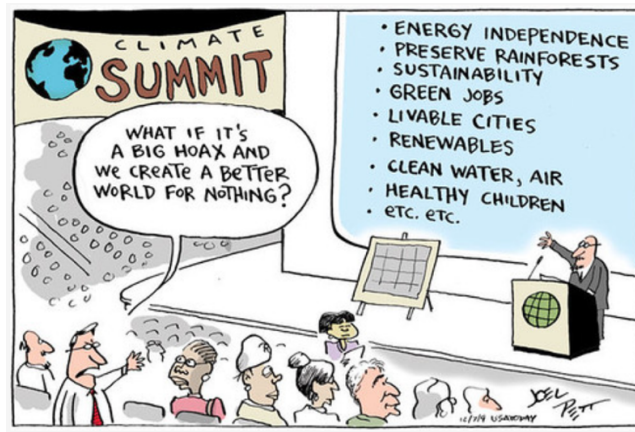


Collage of six images. Image credits below.

Image credits (left to right, top row first):

1. "Solar Panels" by RecondOil is licensed under CC BY 2.0
2. "Climate change – environment" by RecondOil is licensed under CC BY 2.0
3. "MY NEW ENVIRON-MENTALIST FAKE!" by The PIX-JOCKEY (visual fantasist) is licensed under CC BY-NC 2.0
4. "Coal power plant" by eutrophication&hypoxia is licensed under CC BY 2.0
5. "Global Warming (Effetto Serra)" by The PIX-JOCKEY (visual fantasist) is licensed under CC BY-NC 2.0
6. "environment pollution" by RecondOil is licensed under CC BY 2.0

Cartoon for homework



"091207usatoday global warming.91" by RedMonk Red is licensed under CC BY-NC 2.0

Worksheets

Vocabulary list

Here is an example of what a vocabulary list with collocations could look like:

English	German	Collocations	Example sentence
carbon dioxide [kɑːbndaɪ'ɒksaɪd] <i>n</i> , SYNONYM: CO ₂ [siː.əʊ.tuː]	Kohlenstoffdioxid, CO ₂	<ul style="list-style-type: none"> to reduce carbon (dioxide) CO₂ pollution global carbon emissions 	By 2030 the U.S. will reduce its carbon dioxide emissions by nearly 30% below 2005 levels.
emissions [ɪ'mɪʃəns] <i>n</i>	Ausstöße	<ul style="list-style-type: none"> to reduce emissions to cut global CO₂ emissions to control emissions 	The factory promised to reduce the emissions of greenhouse gases.
emission targets/norms/standards etc. [ɪ'mɪʃən'tɑːɡɪts/nɔːrms/'stændəds] <i>n</i>	Emissionsziele, Abgasnormen	<ul style="list-style-type: none"> to meet emission standards to comply with emission norms to set emission limits 	New technologies would help them meet the emission standards .
susceptible to sth [sə'septəbl] <i>adj</i>	anfällig für etw.	<ul style="list-style-type: none"> to be susceptible to pollution to make sb. more susceptible to environmental pollutants 	Children are more susceptible to pollution because they breathe more often, taking in more pollutants.

This is an example of the reading task that could be assigned to each group.

Reading task

Create a table with arguments for and against the motion. Write down the main arguments that the author mentions in the article in the table in 3-5 bullet points on each side. Find out the causes and effects of pollution on people's health and add them to your table. Evaluate how your arguments could support your role and discuss possible strategies to justify your position with your group. Take notes of your results to use for the debate.

Here are links to some example texts you may use. These should be adjusted to your students' needs.

PRO GROUP

- Representative of an environmental organization
<http://redgreenandblue.org/2020/05/10/pollution-kills-9-million-people-year-okay/>
- Young student who supports Fridays for Future
<https://www.renewableenergyworld.com/2016/12/02/what-you-need-to-know-about-air-pollution-its-importance-and-effective-solutions/#gref>

CON GROUP

- CEO of a successful car manufacturer
<https://www.dw.com/en/german-car-lovers-want-to-challenge-fridays-for-future/a-51465631>
- Parent of four who cannot afford an environmentally-friendly lifestyle
<https://www.theguardian.com/lifeandstyle/2019/jul/29/eco-friendly-going-green-poor-cost-diary>

Notes

1. I.e., “an understanding of your influence on climate and climate’s influence on you and society” (Climate.gov 2009)

PART V

LESSON IDEAS: VOCATIONAL EDUCATION AND ENGLISH FOR SPECIFIC PURPOSES (ESP)

Corpus-informed EFL lessons for vocational education and English for Specific Purposes (ESP)

- Developing ESP learning activities for vocational education with a custom virtual corpus by **Cara Symanzik**
- Exploring hospital interactions with Grey's Anatomy by **Simone Campos Silva**
- Creating teaching materials for nursing schools using medical TV series by **Katina Vorbrink** and **Anna Fankhauser**

14. Developing ESP learning activities for vocational education with a custom virtual corpus

Finding and exploiting topic-specific adjectives

CARA SYMANZIK

1 Introduction and rationale

Within the German dual training system, students spend 70% of their study time learning on-the-job at a company and the remaining 30% in publicly funded vocational schools. In vocational schools, teachers of English as a foreign language usually hold English classes in various departments and must thus design relevant learning activities for a diverse student body that could potentially range from future car mechanics to hairdressers-to-be. The three didactical principles “competence development”, “activity orientation”, and “cooperation of learning places” serve as basis for the classes held in vocational schools (Niedersächsisches Kultusministerium 2009: 1-2) and, in English lessons, the four areas of competence – namely receptive skills, productive skills, interaction, and mediation – must be addressed (Niedersächsisches Kultusministerium 2009: 4).

For all four aforementioned areas of competence, students will need to be familiar with vocabulary that is relevant to their future working lives. Consequently, the teacher needs to have specialized knowledge of the kind of idiomatic vocabulary that is used in a particular professional domain. I have observed that existing coursebooks often do not provide genuinely meaningful or relevant vocabulary for specific fields of vocational education in sufficient quantity. This chapter therefore focuses on building a virtual corpus and utilizing it for compiling a list of subject-specific, relevant vocabulary in order to develop language learning activities for specific areas of vocational education.

Outline and objectives

The focus will be on building a virtual corpus and using it for compiling a highly specialized vocabulary list. This vocabulary list is further utilized to develop various language learning activities. In this chapter, the dual apprenticeship as a hairdresser is used illustratively. The following example is localized within the learning area 11 in the third year of training for hairdressers (Kultusministerkonferenz 2008: 19). An exemplary lesson sequence is laid out in Table 1.

Teacher preparation time:

- building the virtual corpus: approx. 15 minutes

- compiling a vocabulary list: approx. 15 minutes
 - designing additional language learning activities: 30 – 60 minutes per activity
- Exemplary language learning activities can be found in Appendices I and II.

Classroom time:

- cloze text: 10 to 15 minutes in class
- guessing card game taboo: 15 to 30 minutes in class

Table 1: Exemplary lesson sequence: Analyzing and describing the skin of a customer

Topic: Designing the skin with decorative cosmetics	
Lesson sequence: Analyzing and describing the skin of a customer	
Contents	Area of competence
Reading a professional text about skin conditions	Receptive skills
Preparing for a role play with the customer	Productive skills
Role play: Analyzing and describing the skin of a customer	Interaction
Telling your colleague about your analysis in German	Mediation

2 Corpus, tools and methods

For building a specialized virtual corpus, the Wikipedia Corpus on the web-based corpus interface www.english-corpora.org/wiki is used. The strength of the Wikipedia Corpus is that it allows to build customized virtual corpora for diverse topics and then conduct searches within these highly specific virtual corpora. This can be used for creating lists of relevant specialized vocabulary. For English teachers in vocational schools, it is highly important to know which vocabulary is genuinely meaningful and relevant for each profession. Crucially, simply looking up a list of words in a dictionary would probably lead to a biased selection of words which the teacher assumes are important, often without direct experience in the students' future field of work. Furthermore, there may not be one-to-one equivalences between the teacher and students' L1 and English. After the virtual corpus has been created, the virtual corpus will be utilized to gather a list of relevant, meaningful and practice-oriented vocabulary. This list will serve as a basis for developing language learning activities which can easily be embedded into a practical use in the classroom.

3 Step-by-step guide

Teacher preparation

Building a virtual corpus

1. Use the browser of your choice to access <https://www.english-corpora.org/>.
2. Go to the tab "my account" and click on "Register". Fill out the registration form and click "SUBMIT". Registration is mandatory and free of charge. You will get an email with a registration link. Confirm

your registration by clicking the link in the email. Go back to the website <https://www.english-corpora.org/> and log in.

3. Click on the tab “home” and then click on the “Wikipedia Corpus”.
4. Click on the tab “Texts/Virtual” and then on the tab “Create corpus” (see Fig. 1).

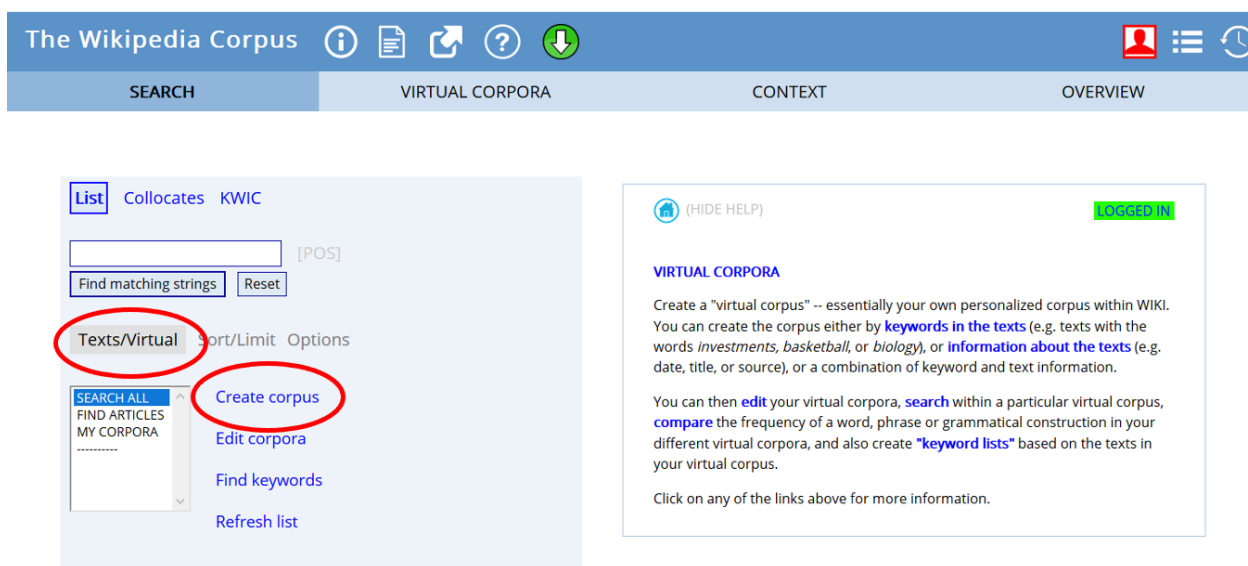


Fig. 1: Building a virtual corpus I.

5. Enter the word *cosmetic* in the field “Title word(s) *”. Choose 100 in the field “# pages” to limit your search to 100 articles. Click “Submit” (see Fig. 2).

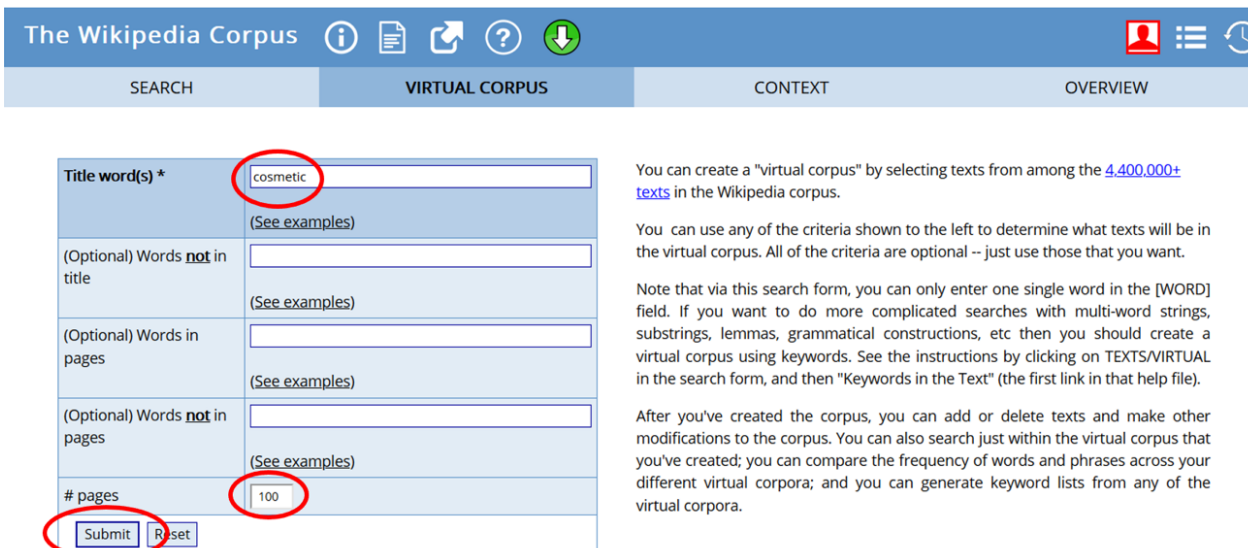


Fig. 2: Building a virtual corpus II.

6. Now you have to look through the results in order to check whether the texts are appropriate to use. You can click on a title to see the full text. You can then select all texts by ticking the empty white box to the left of the “100” and manually untick any texts you wish to exclude. Give your virtual corpus a name, e.g. “SKIN”, by typing it into the “SAVE AS” field and then click on “SUBMIT” to save your virtual corpus (see Fig. 3).

The Wikipedia Corpus

SEARCH VIRTUAL CORPUS CONTEXT OVERVIEW

You can add or remove texts below, and then either save these texts as a new virtual corpus, or else add them to an existing virtual corpus.

SAVE AS: SKIN OR ADD TO: -SELECT- SUBMIT RESET

HELP	<input type="checkbox"/>	# WORDS	↑ ↓	↑ ↓	TITLE (SEE IN WIKIPEDIA)
1	<input checked="" type="checkbox"/>	6,408	894	127	Cosmetics
2	<input checked="" type="checkbox"/>	2,283	137	55	Federal Food, Drug, and Cosmetic Act
3	<input checked="" type="checkbox"/>	759	79	24	MAC Cosmetics
4	<input checked="" type="checkbox"/>	1,900	52	82	Kohl (cosmetics)
5	<input checked="" type="checkbox"/>	1,330	43	48	Jabot Cosmetics
6	<input checked="" type="checkbox"/>	326	42	20	Cosmetic palette
7	<input checked="" type="checkbox"/>	201	27	13	Rouge (cosmetics)
8	<input checked="" type="checkbox"/>	3,008	22	61	Foundation (cosmetics)
9	<input checked="" type="checkbox"/>	1,032	21	43	Fabergé (cosmetics)
10	<input checked="" type="checkbox"/>	937	17	17	Cosmetic dentistry
11	<input checked="" type="checkbox"/>	476	16	17	Compact (cosmetics)
12	<input checked="" type="checkbox"/>	866	14	14	Testing cosmetics on animals
13	<input checked="" type="checkbox"/>	248	13	12	International Nomenclature of Cosmetic Ingredients
14	<input checked="" type="checkbox"/>	531	11	20	NARS Cosmetics
15	<input checked="" type="checkbox"/>	194	10	12	Lakmé Cosmetics
16	<input checked="" type="checkbox"/>	67	9	5	Cosmetic
17	<input checked="" type="checkbox"/>	143	8	1	American Academy of Cosmetic Dentistry

Fig. 3: Screening the results for building a virtual corpus.

- You have now built a virtual corpus on the topic of *skin*. In this example, there are 76 articles with a total of 42,749 words (see Fig. 4).

The Wikipedia Corpus

SEARCH VIRTUAL CORPORA CONTEXT OVERVIEW

MY VIRTUAL CORPORA

HELP	↑ ↓	↑ ↓	LIST NAME	# ARTICLES	# WORDS	FIND KEYWORDS	SPECIFIC	FREQ	CREATED
1	<input type="checkbox"/>	<input type="checkbox"/>	SKIN	76	42,749	NOUN VERB ADJ ADV N+N ADJ+N	<input checked="" type="radio"/>	<input type="radio"/>	0 h

Fig. 4: Virtual corpus for the topic of skin.

Utilizing the virtual corpus for compiling a vocabulary list

- Click on the tab “SEARCH” to use the virtual corpus. You might need to refresh the website and then click on the tab “Texts/Virtual” in order to see your newly built virtual corpus (see Fig. 5).

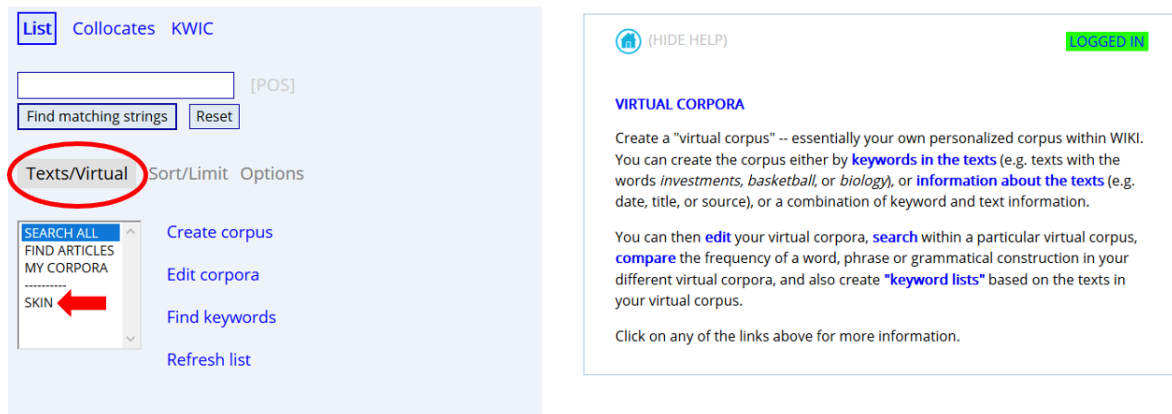


Fig. 5: Selecting your virtual corpus.

- Click on the tab “Collocates”. Enter the word *skin* into the box “Word/phrase”. Click on the field “[POS]” right to the word “Collocates”. The box “Insert PoS” will appear. Click on this box and choose “adj.ALL”. “_j*” appears in the field “Collocates”. Choose the left “1” to find collocates which appear immediately before the target word. Choose the virtual corpus “SKIN” by clicking on it. It is very important that you choose your custom-built virtual corpus, otherwise the whole Wikipedia Corpus will be queried. Finally, click on “Find collocates” (see Fig. 6).

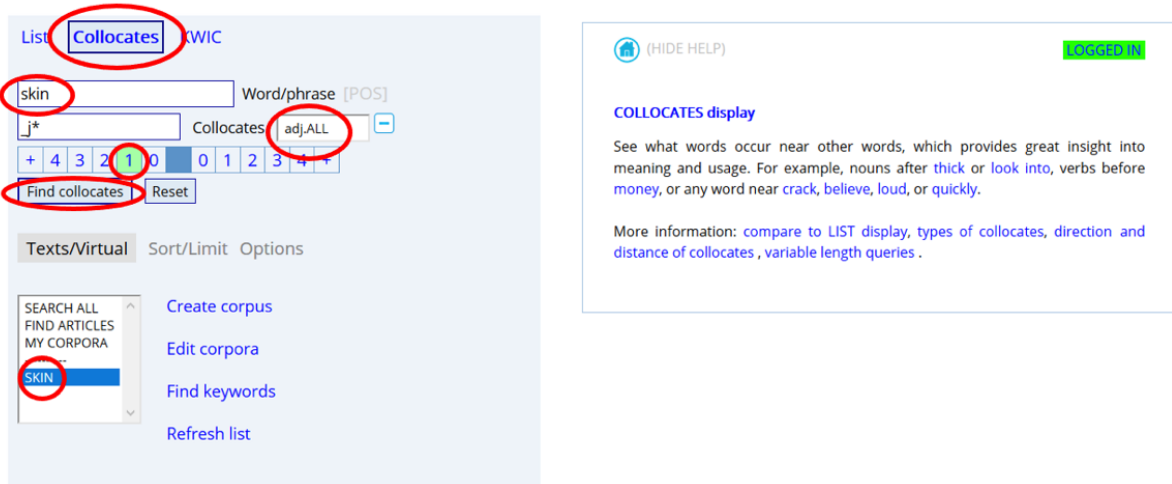


Fig. 6: Finding collocates within your virtual corpus.

- In this example, I obtained a list of 39 collocates sorted by how often they appear in your virtual corpus (see Fig. 7).

The Wikipedia Corpus

SEARCH FREQUENCY CONTEXT OVERVIEW

SEE CONTEXT: CLICK ON WORD OR SELECT WORDS + [CONTEXT] [HELP...]

	<input type="checkbox"/>	CONTEXT	FREQ
1	<input type="checkbox"/>	OILY	9
2	<input type="checkbox"/>	PALE	6
3	<input type="checkbox"/>	DRY	6
4	<input type="checkbox"/>	NORMAL	5
5	<input type="checkbox"/>	SENSITIVE	5
6	<input type="checkbox"/>	UNEVEN	4
7	<input type="checkbox"/>	ACNE-PRONE	3
8	<input type="checkbox"/>	WHITE	2
9	<input type="checkbox"/>	DEAD	2
10	<input type="checkbox"/>	DARK	2
11	<input type="checkbox"/>	DAILY	2
12	<input type="checkbox"/>	COMMON	2
13	<input type="checkbox"/>	OLD	2
14	<input type="checkbox"/>	TATTOOED	1
15	<input type="checkbox"/>	TANNED	1
16	<input type="checkbox"/>	SURROUNDING	1
17	<input type="checkbox"/>	SMOOTH	1
18	<input type="checkbox"/>	ORGANIC	1
19	<input type="checkbox"/>	ROUGH	1
20	<input type="checkbox"/>	RIGHT	1
21	<input type="checkbox"/>	RICHER	1

Fig. 7: Collocates for the word skin filtered by frequency.

4. You now need to filter the words in order to just get relevant words that describe skin. You can see the context of the word by clicking on it. Here, the first word which seems inappropriate is *dead*. By clicking on it, you can see that it is not used for describing skin, but rather *skin cells*. Filter out such words not describing skin itself (see Fig. 8).

The Wikipedia Corpus

SEARCH FREQUENCY CONTEXT OVERVIEW

(SHUFFLE)

CLICK FOR MORE CONTEXT [?] SAVE LIST CHOOSE LIST [-----] CREATE NEW LIST [] [?]

1	Cosmetics	A	B	C	Asia. # #Exfoliants# # Exfoliants are products that help slough off dry, dead skin cells to improve the skin's appearance. This is achieved either by using acids
2	Cosmetics	A	B	C	circular motion. Gels, creams or lotions may contain an acid to encourage dead skin cells to loosen, and an abrasive such as beads, sea salt, sugar

Fig. 8: Words in context.

5. In this example, the following words are filtered out: *dead, daily, common, surrounding, organic, right, richer, related, rejuvenate, natural, individual, improving, general, exotic, and basic*. In order to get the most meaningful words, you might also delete the words *old*, and *great* from your list because these are very general and thus do not need to be taught specifically. You may also want to exclude the word *laggy* from the list, because it is a colloquial word, which is not used to describe skin professionally. Next, make a list of the adjectives you have selected. You may also add translations (see Table 2).

Table 2: Exemplary vocabulary list compiled using the virtual corpus.

Number	Word (English)	Word (German)
1	oily	fettig
2	pale	bläss
3	dry	trocken
4	normal	normal
5	sensitive	empfindlich
6	uneven	uneben
7	acne-prone	zu Akne neigend
8	white	weiß
9	dark	dunkel
10	tattooed	tätowiert
11	tanned	gebräunt
12	smooth	weich
13	rough	rau
14	reddened	gerötet
15	problematic	problematisch
16	moisturised	hydratisiert
17	mature	reif
18	flaky	schuppig
19	blemish-prone	zu Unreinheiten neigend
20	ashy	aschfahl
21	flawless	makellos

Preparing teaching materials for language learning activities

Students need to know how to use newly-learnt words in written and spoken language correctly in order to utilize them holistically. They further should be able to use the new words in all four areas of competence. Teachers should make sure that students get the opportunity to learn the new words and that they get the chance to practise the pronunciation as well as the spelling. The compiled vocabulary list allows teachers to develop a range of language learning activities. In this chapter, I demonstrate how to use the list in Table 2 to create a controlled language learning activity, in the form of a cloze text, as well as a communicative language learning activity, in the form of a taboo card game.

Gap-filling activities can be used to familiarize students with both the meaning and spelling of new words. The teacher needs to prepare a cloze text. Depending on the level of the individual students, the words which should be filled in the gaps can be displayed below the text as an assistance or the first letter of each

word can be given in the gaps. Differentiation in mixed-ability classes is thus made possible. An example of a cloze text – with an attached sample solution – can be found in Appendix I. Whilst developing a cloze text, sentences taken directly from the corpus can be used in order to ensure that authentic examples are demonstrated.

The vocabulary list compiled from the virtual corpus can also be used to create taboo playing cards. This is a game, which encourages the students to communicate and use new vocabulary in spoken language. This method can be used after the new words have been introduced and then in different stages of the lesson sequence in order to revise the vocabulary. Exemplary taboo playing cards can be found in Appendix II.

4 Options and further ideas

Students in vocational education need to know subject-specific adjectives in order to describe different words in greater detail. Students must know meaningful adjectives for varying situations and a virtual corpus is an appropriate instrument for creating highly specific vocabulary lists in the field of vocational education. The outlined procedure and the exemplary language learning activities of this chapter can be immediately used for the dual apprenticeship as a hairdresser. The exemplary topic can also straightforwardly be utilized for the dual apprenticeship as a beautician/cosmetician.

This chapter demonstrates that general adjectives are not sufficient for describing specific words. This also applies for other professional disciplines in the field of vocational education, for example metal work, wood work, electrical engineering, glass engineering, home economics, social education, healthcare professions, and many others. The procedure for building and utilizing a virtual corpus can easily be adapted and used for different types of schools within the field of vocational education. The vocabulary lists can be further expanded based on the proficiency level of the learners. Also, the virtual corpus can be used in order to compile vocabulary lists of different word classes, for example subject-specific nouns, verbs, and other collocations (e.g. verb-noun collocations).

5 Caveats and limitations

Since the Wikipedia Corpus is a compilation of written articles, it may not always be adequate for teaching spoken language. Nevertheless, the Wikipedia Corpus can provide a useful, general overview of the kind of words and collocations which are relevant for, and frequently used in specific areas.

Note that whilst using the website <https://www.english-corpora.org/>, from time to time a message will appear telling you to upgrade your account. You can just wait for a few seconds and then click on “Click here to continue with your search”. The message will disappear and you do not need to upgrade your account (see Fig. 9).

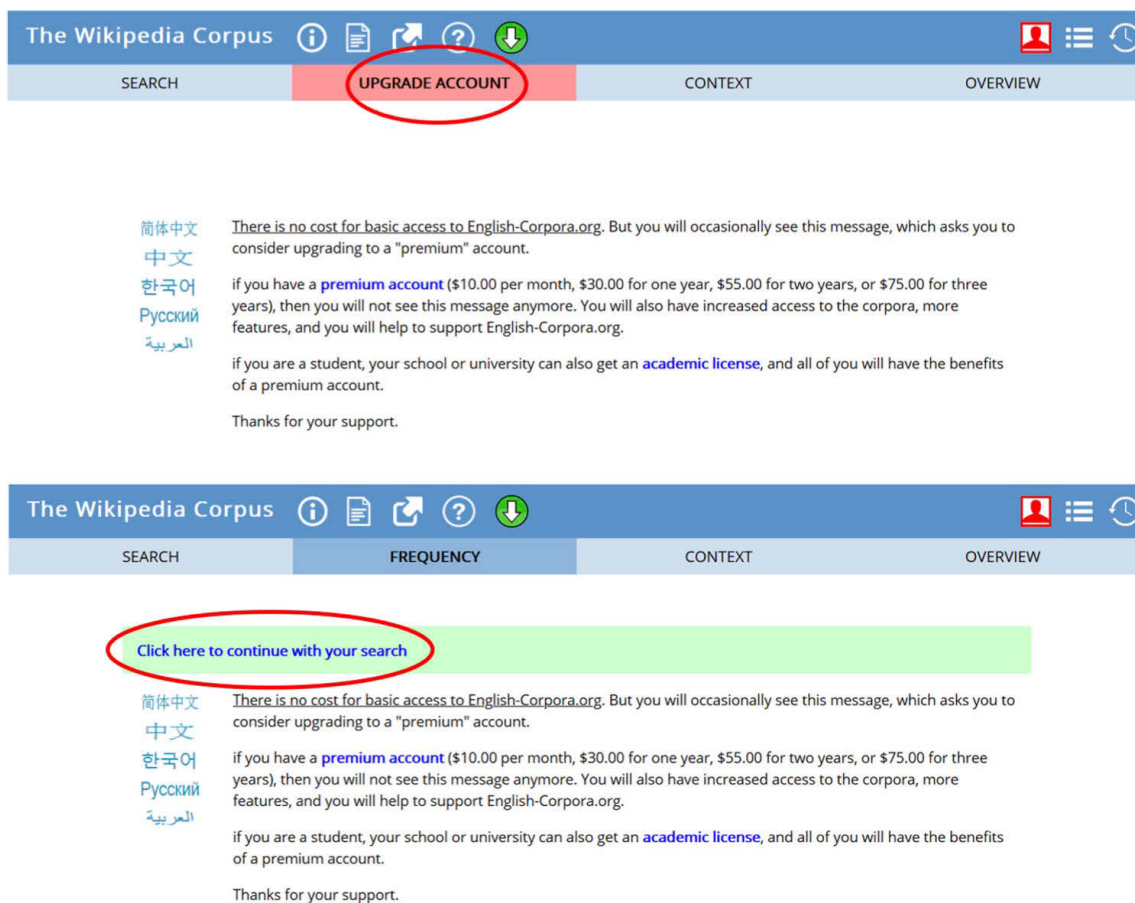


Fig. 9: Message to upgrade the account

6 Conclusion

In this chapter, I explained how to build a virtual corpus and utilize it for developing language learning activities for specific areas of vocational education. The three didactical principles as well as the four areas of competence relevant to vocational education – especially in the dual apprenticeship model – were taken into account. After building the virtual corpus, I explained how to use it to compile a subject-specific vocabulary list. This vocabulary list served as a basis for developing two language learning activities. These activities can be immediately embedded into teaching English in classes for hairdressers as well as beauticians / cosmeticians. I further explained how the aforementioned procedure can be adapted for different types of schools and various professional disciplines within vocational education. In addition, I mentioned options for adapting the activities to heterogeneous levels of proficiency. In conclusion, the procedure of building a virtual corpus for a specific area of vocational education in order to compile a vocabulary list to develop teaching materials has proven itself to be user-friendly and broadly applicable for language teachers in the field of vocational education.

7 Resources and references

Kultusministerkonferenz. 2008. Rahmenlehrplan für den Ausbildungsberuf Friseur/Friseurin: Beschluss der Kultusministerkonferenz vom 10.04.2008. <https://www.kmk.org/fileadmin/pdf/Bildung/BeruflicheBildung/rlp/Friseur.pdf> (14 October, 2020).

Niedersächsisches Kultusministerium. 2009. Rahmenrichtlinien für das Unterrichtsfach Englisch/ Kommunikation in der Berufsschule. https://www.nibis.de/nli1/bbs/archiv/rahmenrichtlinien/eng_bs_an.pdf (14 October, 2020).

Wikipedia Corpus on english-corpora.org. <https://www.english-corpora.org/wiki/> (14 January, 2021).

8 Appendices

Appendix I: Cloze text

Describing the skin of customers with adjectives

Hairdressers and cosmeticians need to know about their customers' skin in order to choose cosmetic treatments. Some customers have great skin. This skin has no anomalies at all and is called _____ skin. This skin is _____ and _____.

Some customers have _____ skin – especially in winter when the sun is not shining that much. The color of the skin is nearly _____. In summer, some customers have _____ and _____ skin.

Some customers also like to decorate their skin with permanent colors. They have _____ skin. Older customers have special needs. These customers have _____ skin which needs a special treatment. The skin of older customers can also look dehydrated and discolored and is then called _____ skin.

Some customers do not have good skin; they have very _____ skin. This skin tends to be tender and is called _____ skin. The skin can be _____, and when it does not produce enough oil and lacks moisture. When the skin starts producing too much oil, the skin can become _____. Too much oil can lead to pimples which can result in customers having _____ skin. This skin is often _____ and can even become _____ when specific bacteria is on the skin. This skin is often irritated and is called _____ skin.

It is very important for hairdressers and cosmeticians to know that!

Fill in the following words:

dry, normal, sensitive, pale, oily, uneven, white, acne-prone, dark, problematic, reddened, smooth, rough, tanned, tattooed, ashy, blemish-prone, mature, flaky, moisturized

Sample solution

Hairdressers and cosmeticians need to know about their customers' skin in order to choose cosmetic treatments. Some customers have great skin. This skin has no anomalies at all and is called **normal** skin. This skin is **smooth** and **moisturized**.

Some customers have **pale** skin – especially in winter when the sun is not shining that much. The color of the skin is nearly **white**. In summer, some customers have **dark** and **tanned** skin.

Some customers also like to decorate their skin with permanent colors. They have **tattooed** skin. Older customers have special needs. These customers have **matured** skin which needs a special treatment. The skin of older customers can also look dehydrated and discolored and is then called **ashy** skin.

Some customers do not have good skin; they have very **problematic** skin. This skin tends to be tender and is called **sensitive** skin. The skin can be **dry, flaky** and **rough** when the skin does not produce enough oil and lacks moisture. When the skin starts producing too much oil, the skin can become **oily**. Too much oil can lead to pimples which can result in customers having **uneven** skin. This skin is often **blemish-prone** and can even become **acne-prone** when specific bacteria is on the skin. This skin is often irritated and is called **reddened** skin. It is very important for hairdressers and cosmeticians to know that!

Appendix II: Taboo

<p>The customer has dry skin.</p> <p>flaky moisture oil</p>	<p>The customer has normal skin.</p> <p>basic perfect imperfections</p>	<p>The customer has sensitive skin.</p> <p>tender careful problematic</p>	<p>The customer has white skin.</p> <p>pale fair snow</p>	<p>The customer has dark skin.</p> <p>black tanned sun</p>
<p>The customer has pale skin.</p> <p>white fair snow</p>	<p>The customer has oily skin.</p> <p>oil fat grease</p>	<p>The customer has uneven skin.</p> <p>imperfection pimple rough</p>	<p>The customer has ashy skin.</p> <p>old mature discolored</p>	<p>The customer has tattooed skin.</p> <p>color permanent art</p>
<p>The customer has moisturized skin.</p> <p>normal hydrated oil</p>	<p>The customer has flaky skin.</p> <p>dry rough moisture</p>	<p>The customer has problematic skin.</p> <p>blemish pimple dry</p>	<p>The customer has reddened skin.</p> <p>irritated rash sun</p>	<p>The customer has smooth skin.</p> <p>moisture oil soft</p>

1 moisturised_{BE}/moisturized_{AE}. Note that the corpus featured the British English spelling, but I would recommend familiarizing students with both spelling variants.

15. Exploring hospital interactions with Grey's Anatomy

Building and exploiting a TV series corpus for nursing vocational schools

SIMONE CAMPOS SILVA

1 Introduction and rationale

“Words – so innocent and powerless as they are, as standing in a dictionary, how potent for good and evil they become in the hands of one who knows how to combine them.” (Hawthorne)

The International Council of Nurses states:

Nursing encompasses autonomous and collaborative care of individuals of all ages, families, groups and communities, sick or well and in all settings. Nursing includes the promotion of health, prevention of illness, and the care of ill, disabled and dying people. Advocacy, promotion of a safe environment, research, participation in shaping health policy and in patient and health systems management, and education are also key nursing roles. (ICN, 2002)

Communication and interaction are particularly important competences to be able to fulfil the tasks described above. This year's newly designed and revised German nursing apprenticeship introduces a new curriculum, which describes communication as a key nursing activity (Ammende et al. 2019: 51). Moreover, subjects of general education should be taught with connection to nursing topics (Niedersächsische Landesschulbehörde 2019). As English is one of those general education subjects, communication in hospital settings can be efficiently combined with learning English. Talking to and supporting patients and their next of kin in difficult and life-threatening situations can be a challenging task during a nursing apprenticeship. Therefore, practicing those kinds of situations in a protected and sheltered environment is of crucial importance. The series *Grey's Anatomy* offers the possibility to examine many different challenging interactions and communications in hospital settings. Consequently, students have the opportunity to learn from those interactions and develop communication strategies in the most widely used lingua franca, English.

Outline and objectives

- This lesson is designed for B1 to C1 vocational school learners of English (intermediate to advanced) who are training to become nurses.
- The aim is to focus on communication and interaction in hospital settings in order to support and help nursing students.
- Teachers should plan around 2-2.5 hours to follow all steps described to prepare the class

session.

2 Corpus, tools and methods

The corpus building tool Sketch Engine is used for this chapter. Sketch Engine offers the possibility to build and analyse custom-made corpora compiled from texts from the web. Unfortunately, Sketch Engine is not free to use, but there is a free 30-day trial, which is more than sufficient for this lesson. Simply create a free trial account and you will be able to start working with corpora.

Grey's Anatomy was created by Shonda Rhimes and is a “prime-time American television medical drama that debuted on the American Broadcasting Company (ABC) network in 2005. [...] The program focuses on the personal and professional lives of surgical interns and their medical mentors.” (Encyclopaedia Britannica 2018). Momentarily, *Grey's Anatomy* comprises 16 seasons. Due to its award-winning success, the series is famous worldwide. Therefore, students might enjoy learning English with the help of a TV drama they know.

In this chapter, Sketch Engine's corpus building functions and the *Grey's Anatomy* scripts will be combined to create an English learning environment for nursing students.

3 Step-by-step guide for teachers

Below you will find instructions on how to build your own *Grey's Anatomy* corpus and how you use tools on Sketch Engine to work with this corpus in order to create work sheets for students.

Teacher preparation

Corpus compilation

1. Open the website Sketch Engine (<https://www.sketchengine.eu>). Create an account and then log in.
2. As soon as you are logged into your account, you will find an option to create your own corpus in the upper right corner of your dashboard. Click on “NEW CORPUS” (Fig. 1).

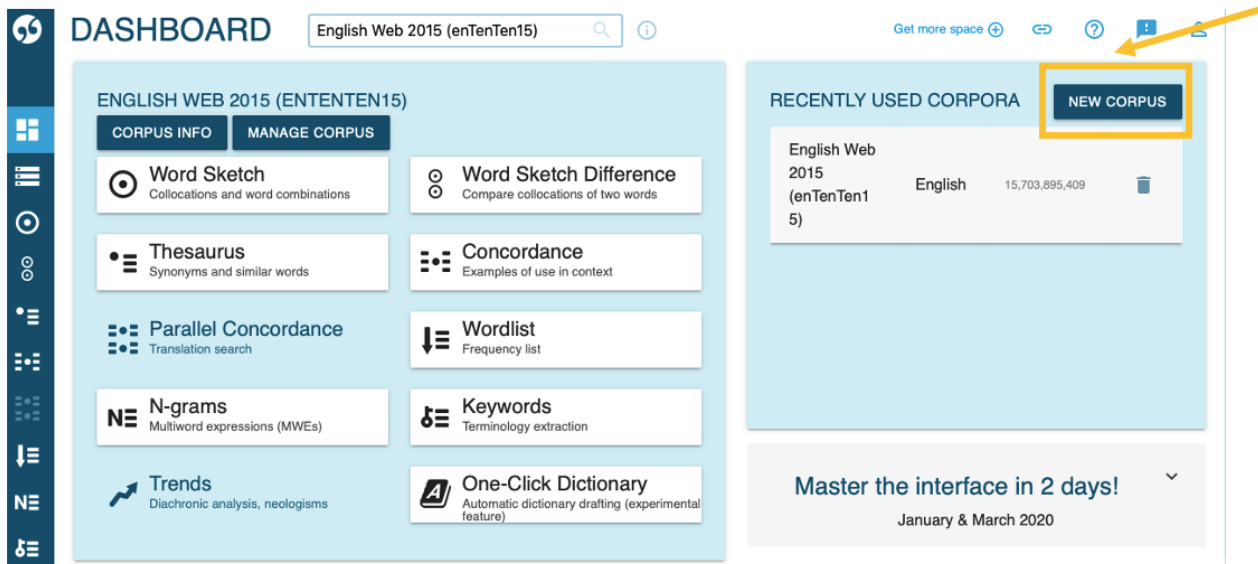


Fig. 1: Create a new corpus on Sketch Engine.

3. In the next step, you will be asked to name your corpus (e.g. *Grey's Anatomy*), to choose a corpus type (single language corpus), to select a language (English), and to shortly describe your corpus if needed (*Grey's Anatomy* scripts). After that, click on the red button "NEXT".
4. Now, you can start adding texts to your newly created corpus. There are two options to do so: you can either find texts on the web or upload your own texts from your computer. In this case, choose "Find texts on the web" (Fig. 2).

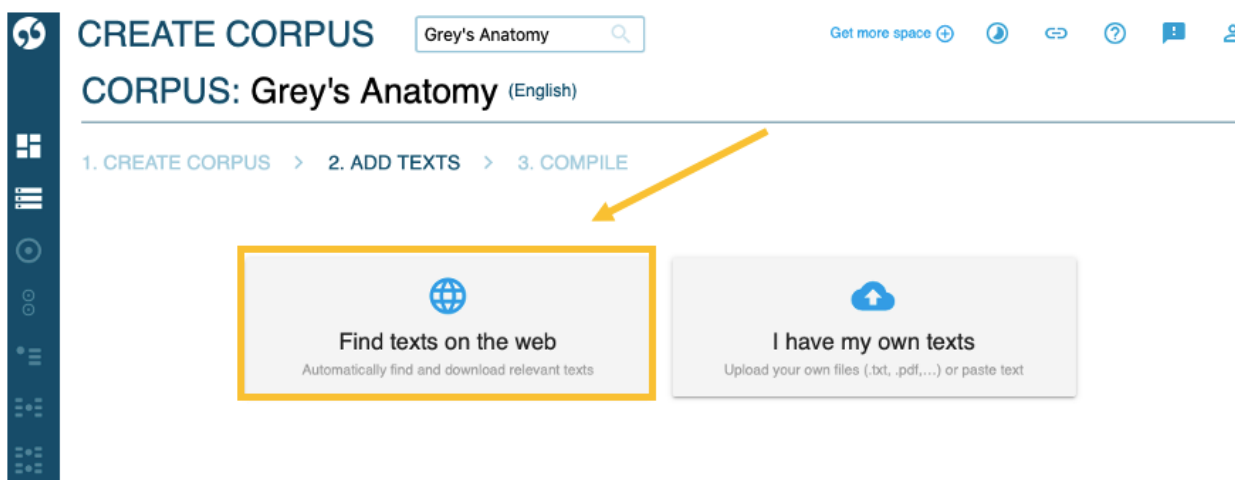


Fig. 2: Options for adding texts to the corpus on Sketch Engine.

5. Before being able to upload texts from the web into your corpus, you need to find a website from which you would like to upload data. In this case, just google "*Grey's Anatomy* scripts" and you will find different pages that offer scripts of all the *Grey's Anatomy* seasons. Choose one of them, e.g.: <https://transcripts.foreverdreaming.org/viewforum.php?f=11> (last accessed 16 June 2020). Unfortunately, it is not possible to just copy the link to the whole website as it – like many other websites – does not only contain *Grey's Anatomy* scripts but also other TV seasons scripts. In order to

make sure that you only upload *Grey's Anatomy* scripts to your corpus, you need to open each episode's webpage separately and copy the URLs. If you decide to choose a different website, just make sure to find one without too many fan comments etc. as all those comments will also be added to your corpus.

6. *Grey's Anatomy* comprises 16 seasons with often more than 20 episodes each. Sketch Engine's default maximum amount of 1,000,000 words for a custom-made corpus will not be enough for the scripts from all 16 seasons. Consequently, you can only upload parts of the *Grey's Anatomy* seasons. 1,000,000 words should enable you to upload about twelve seasons to your corpus.
7. Sketch Engine gives you three different options for finding texts on the web: "Web search", "URLs" or "Website". This time choose the second option: "URLs". Sketch Engine requires you to paste every single URL from every episode you would like to add to your corpus. After pasting one URL, always press "Enter" on your keyboard in order to add the next URL (Fig. 3).
8. It will be easier to manage your corpus later if you create separate folders for each season. To do so, choose folder names: For season 1 = season1, season 2 = season2, season 3 = season3, etc. (Fig. 3).
9. Click on the red button "GO" and Sketch Engine will start to upload data from the webpages and build your *Grey's Anatomy* corpus (Fig. 3).

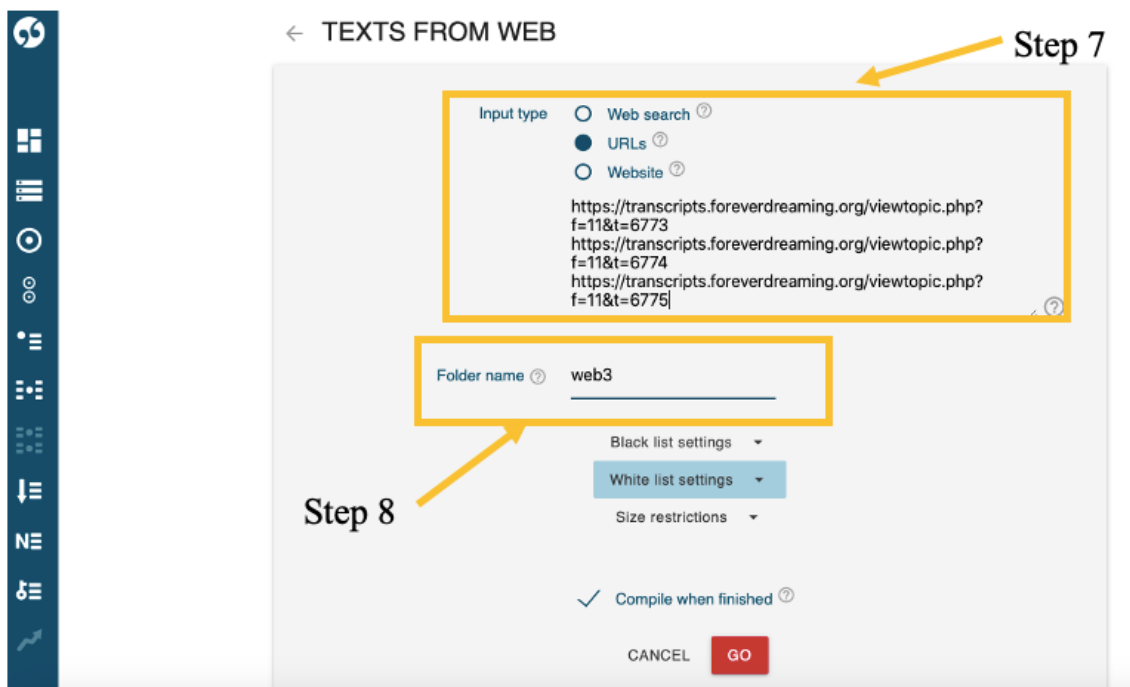


Fig. 3: Paste the URLs of each episode into the corresponding fields.

10. Once the uploading process is finished, you can repeat the exact same procedure with the next season. Therefore, come back to your dashboard (Fig. 1) and click on "Manage corpus" on the upper left side. Then choose the option "Make bigger" (Fig. 4), which will bring you back to the screen in Fig. 2. Now, follow steps 7 to 9 again for as many seasons as you would like to include or until you reach Sketch Engine's maximum free word count of 1,000,000 words.

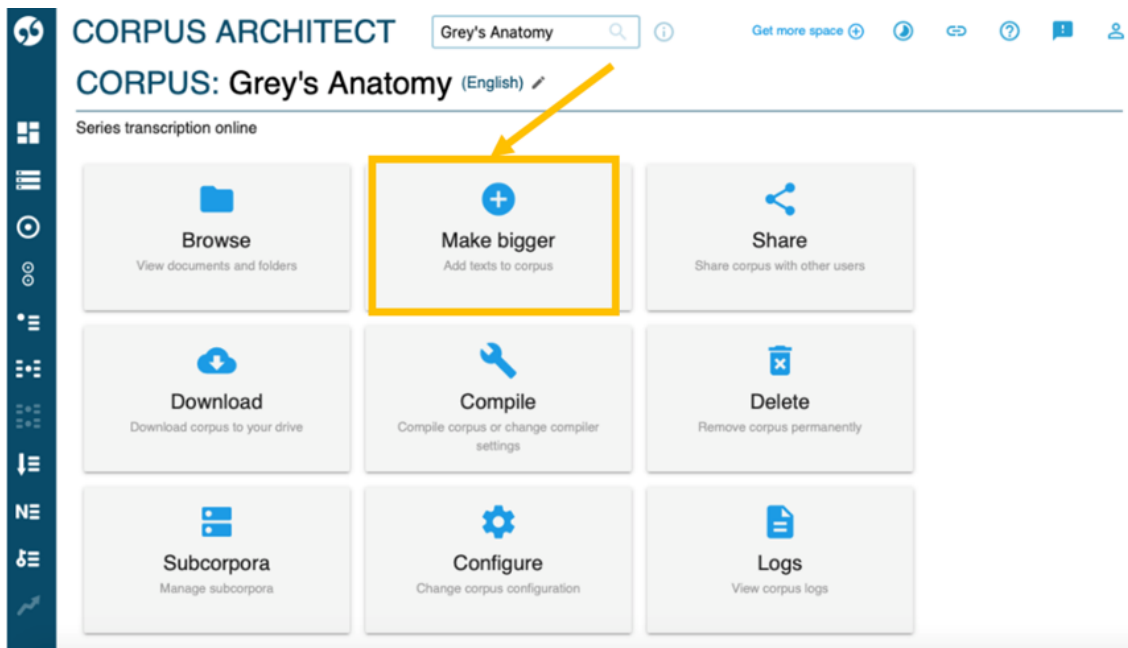


Fig. 4: Options for managing your corpus on Sketch Engine.

Corpus queries and work sheet creation

I will present two functions of Sketch Engine that allow you to exploit your corpus for English language teaching: Word Sketch and Concordance. These functions can be used to create work sheets for your students.

Using 'Word Sketch' to create a vocabulary list on a certain topic

With the help of your newly built *Grey's Anatomy* corpus, it is very easy to build a vocabulary list on a specific topic. You are completely free to decide on a topic. As an example, this chapter will focus on the topic of *organ donation*. Here we shall create a word sketch for the noun *organ*:

1. On your Sketch Engine dashboard, click on the option "Word Sketch" (Fig. 5).

DASHBOARD

Grey's Anatomy

GREY'S ANATOMY

- Word Sketch**
Collocations and word combinations
- Thesaurus**
Synonyms and similar words
- Parallel Concordance**
Translation search
- N-grams**
Multiword expressions (MWEs)
- Trends**
Diachronic analysis, neologisms
- Word Sketch Difference**
Compare collocations of two words
- Concordance**
Examples of use in context
- Wordlist**
Frequency list
- Keywords**
Terminology extraction
- One-Click Dictionary**
Automatic dictionary drafting (experimental feature)

Fig. 5: Dashboard on Sketch Engine to work with the corpus.

2. Use the basic search function. You will be asked to type in a lemma. In this case, type in *organ* and click on "GO" (Fig. 6).

WORD SKETCH

Grey's Anatomy

Get more space +

BASIC | ADVANCED | AS A LI... | ABO...

Search ?

organ

GO

Fig. 6: Word Sketch search on Sketch Engine.

3. Finally, you will obtain lists of words that are frequently associated with the word *organ* in *Grey's Anatomy* (Fig. 7).

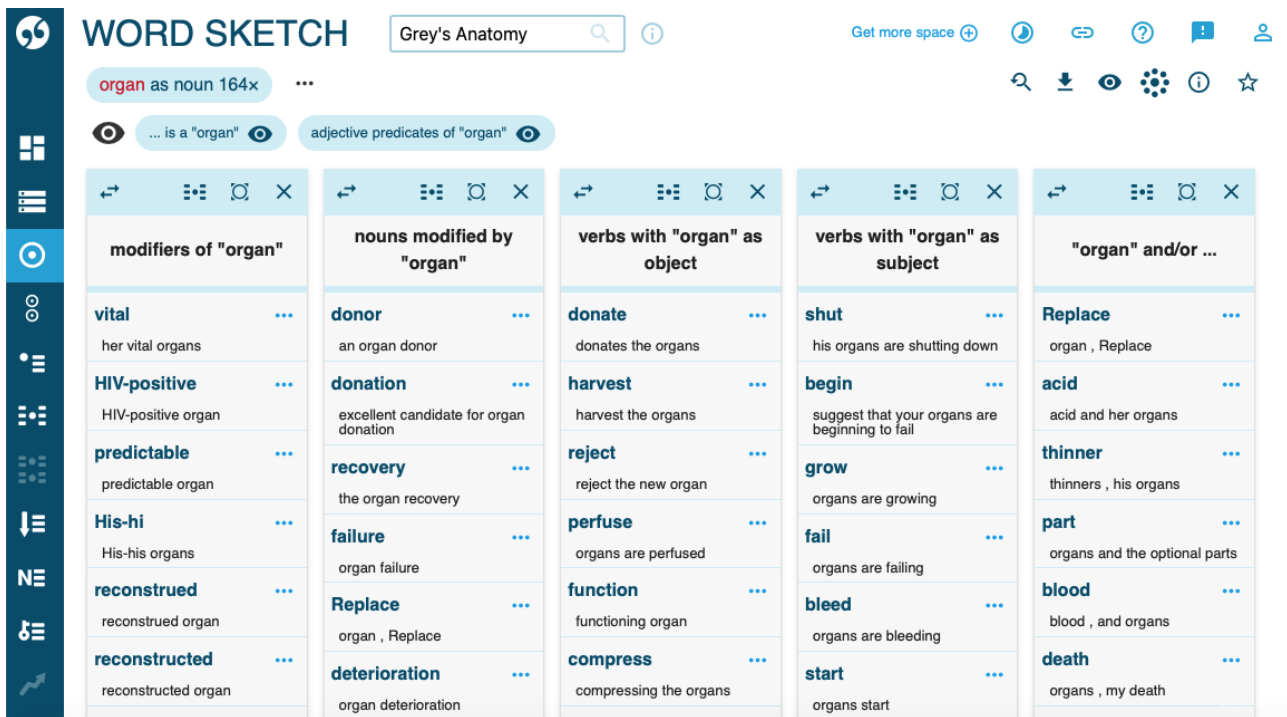


Fig. 7: Word Sketch for organ.

- With the help of your Word Sketch of the word *organ*, you can create a vocabulary list including useful collocations and prepositions, for your students. Copy the words and collocations that you want your students to learn and paste them in a separate document. Consider choosing vocabulary based on your pedagogical experience. As the Word Sketch shows the most frequent collocates, it is up to you to choose useful vocabulary. An example vocabulary list can be found in Appendix 1.
- Students can practise the new vocabulary by filling in a cloze text based on the Word Sketch list. To create such a cloze exercise, click on “Concordance – see all collocates in this column in context” (Fig. 8).

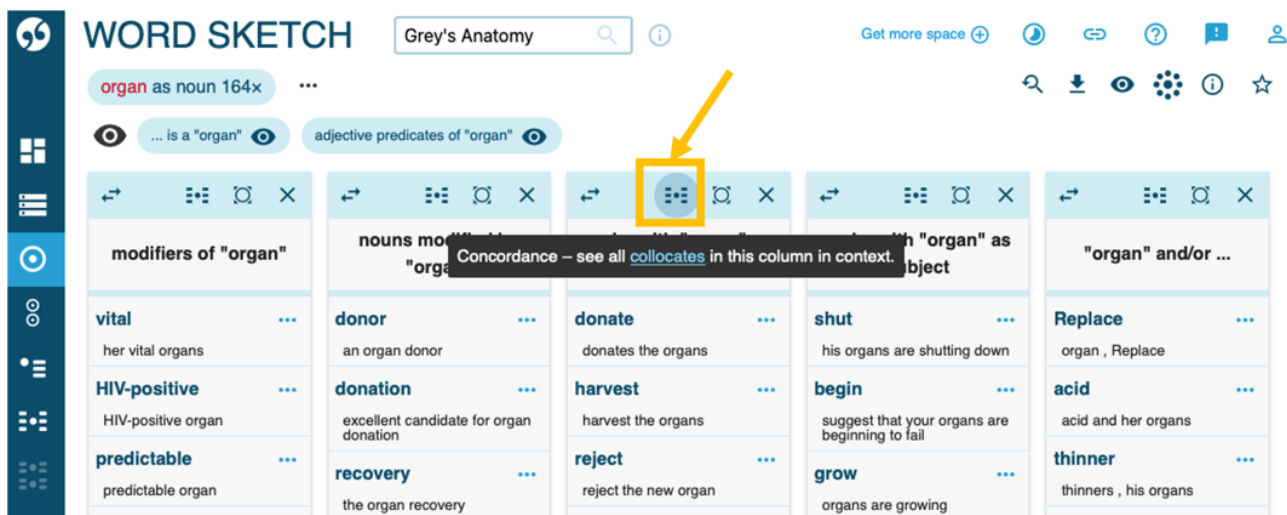


Fig. 8: Finding the context to the words in the Word Sketch.

- In the menu, click on “KWIC” and select the option “sentence” (Fig. 9). You will see how the words from your vocabulary list are used in context of the *Grey's Anatomy* scripts.

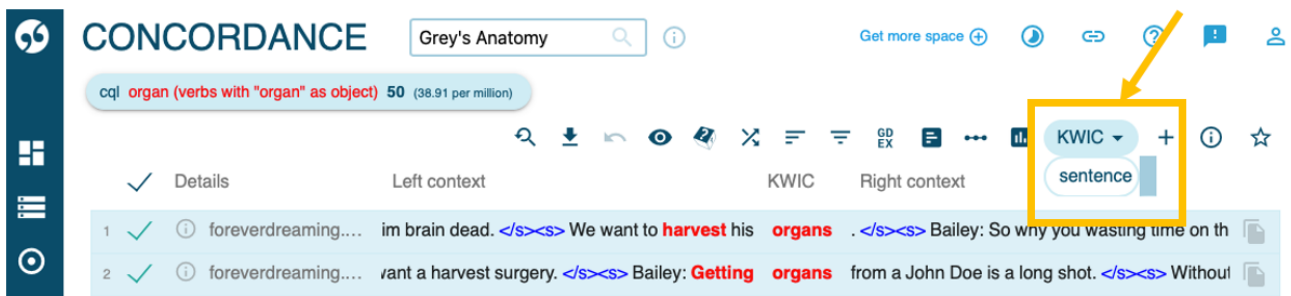


Fig. 9: Sentences from the Word Sketch in a Concordance list.

7. Decide on which sentences you want to add to your cloze exercise, copy them and add appropriate gaps. An example cloze exercise can be found in Appendix 2.

Using 'Concordances' to extract dialogues from the episodes

Nursing students may worry about how to address patients and their next of kin about difficult situations such as a life-threatening diagnosis, death or organ donation. Staff-patient-interactions in *Grey's Anatomy* provide a very valuable source of such interactions in English. Needless to say that such constructed conversations and interactions cannot exactly mirror real hospital situations. Nevertheless, they can be used to practise dealing with emotional situations in a safe classroom environment. Students may benefit from these dialogues by focusing on discourse markers, fillers, distance and empathy, or other communicative strategies in dialogues.

Again, the *Grey's Anatomy* corpus proves to be a useful tool in finding the interactions mentioned above:

1. Go to the dashboard. Click on "Concordance" (Fig. 10).

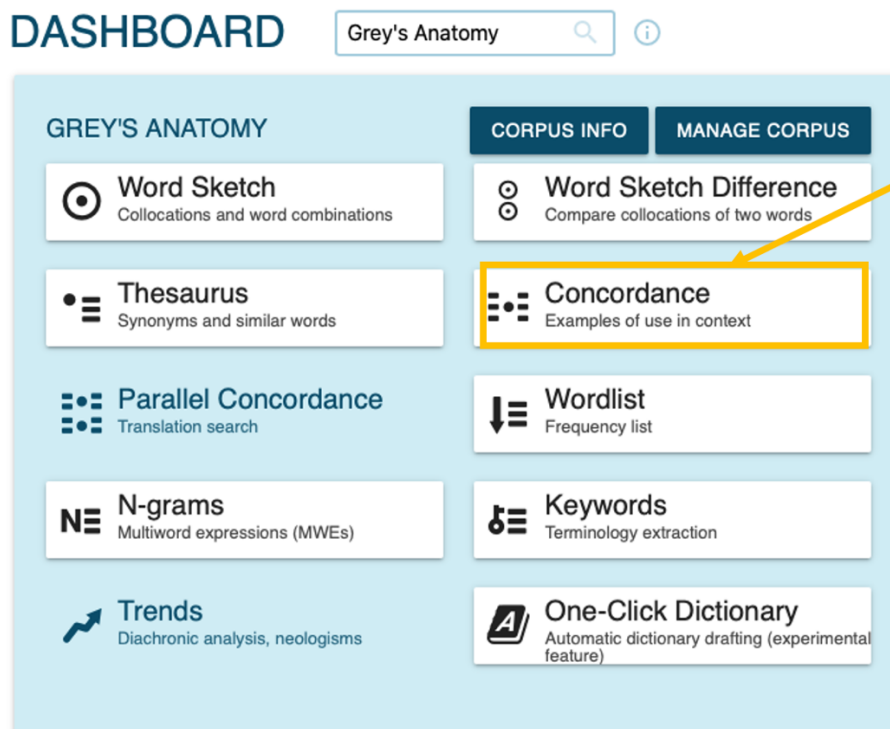


Fig. 10: Select the 'Concordance' function on the dashboard screen.

- Use the basic search function and type in *organ don** in order to find text passages in your corpus that include the words *organ donor/donors/donation*. Further possible results might be *organ donating/organ don't* etc. Click on the red button "SEARCH" (Fig. 11).

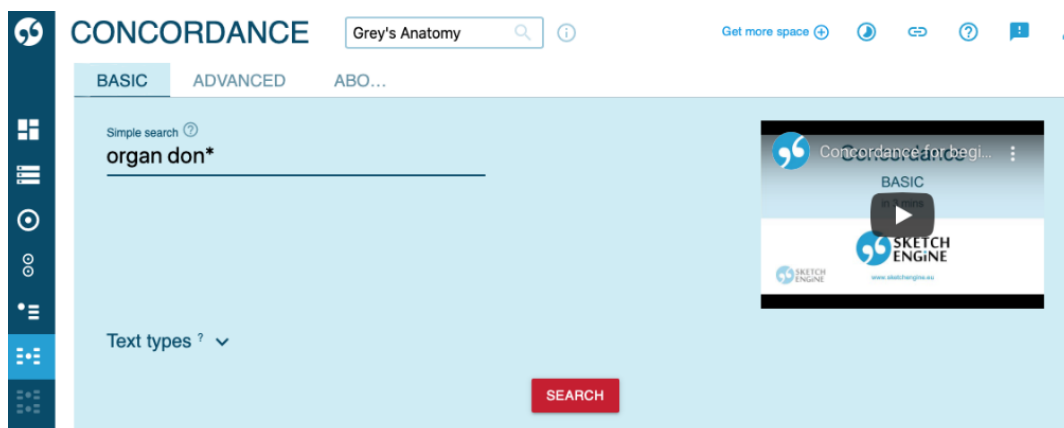


Fig. 11: Basic search option function to extract concordances.

- Sketch Engine will show you all the text passages in your corpus that include the words *organ donor/donors/donation* (Fig. 12).

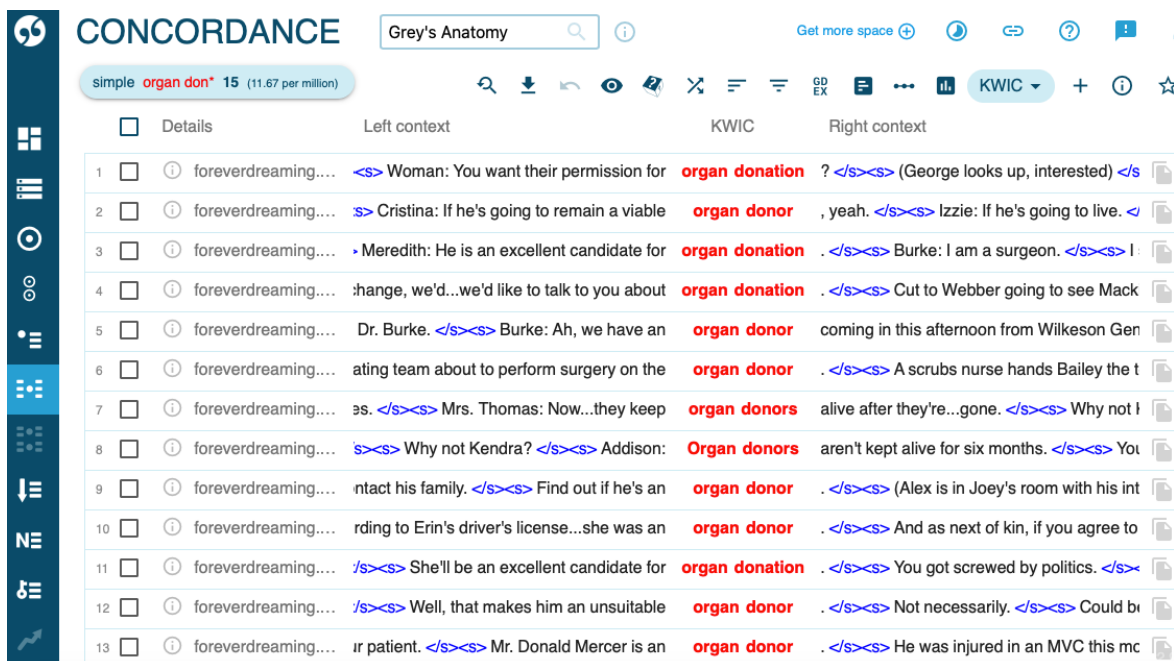


Fig. 12: Concordance list of the words *organ don**.

- Choose one of the given text passages. clicking on the red writing (KWIC), you can read more from the scene about organ donation which might be helpful to decide on a suitable text passage (Fig. 13).

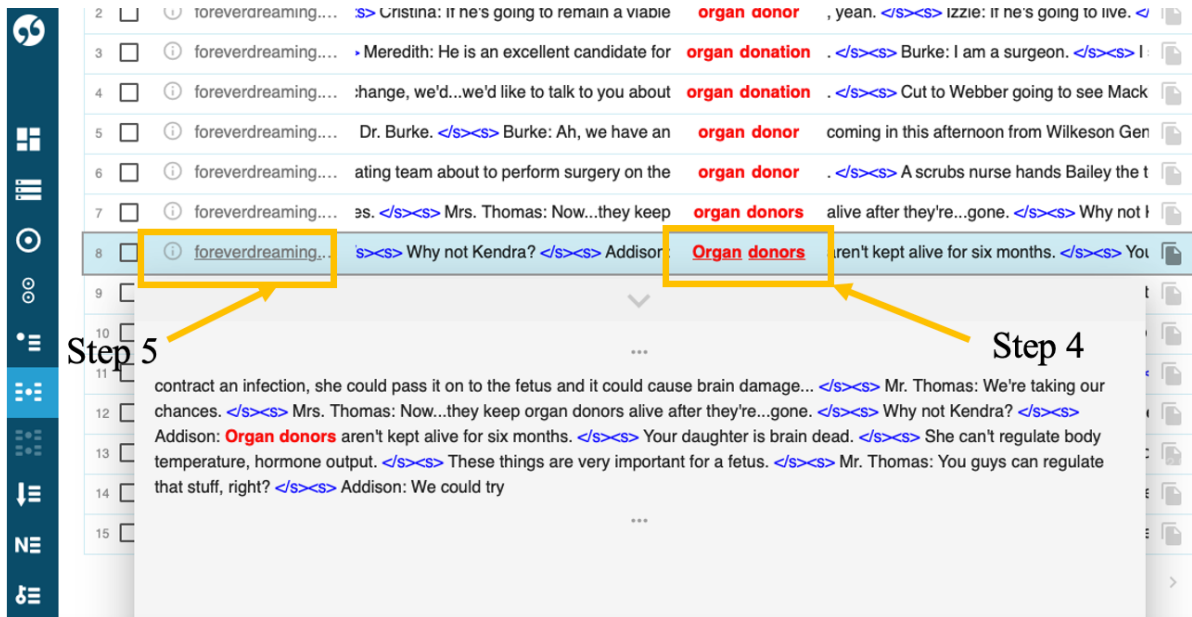


Fig. 13: Longer extract of the script in the Concordance list.

- After having found a suitable text passage, click on the information icon or the URL next to the concordance line to obtain more information about the episode in which this scene can be found. Click on the URL and you will be forwarded to the full *Grey's Anatomy* script (Fig. 14).



Fig. 14: Details about the text passages in the Concordance list.

- Use the search function in your browser to find the text passage in the full *Grey's Anatomy* script displayed on that web page and copy the relevant text passage from the script. An example of a dialogue on organ donation based on the series *Grey's Anatomy* and found with this method can be found in Appendix 3.
- With the help of the Concordance list on Sketch Engine it is possible to find suitable text passages in the *Grey's Anatomy* scripts where medical staff talk about organ donation to the next of kin. If possible, you could also show the video of that scene as an introduction to the topic.

Lesson plan and tasks

Table 1: Lesson Plan for 90 min.

Time/Phase	Content	Method/Media
5 minutes	Show sequence from <i>Grey's Anatomy</i> episode about organ donation and doctors talking to next of kin	Clip from <i>Grey's Anatomy</i> episode
10 minutes	Defining the term <i>organ donation</i> .	Classroom discussion; Blackboard
20 minutes	Students receive a cloze text based on the Word Sketch and Concordance lines from Sketch Engine, which they can fill in with help of the vocabulary list. Compare results.	Cloze Vocabulary list
15 minutes	Students receive three different dialogues about organ donation illustrating staff-patient-interactions. Tasks: How is this interaction structured? How well did the participants handle the situation? What could be improved?	Dialogues from the series
20 minutes	Students receive a new task: Imagine taking care of a brain-dead patient. How would you address the possibility of organ donation to the next of kin? Remember the introduction sequence. Develop a dialogue with a partner.	Pair work
10 minutes	Students present their dialogues (voluntarily) and their peers give feedback.	-
10 minutes	Create a list of phrases students might want to remember for staff-patient-interactions in general.	-

4 Options and further ideas

The class session described above focuses on the topic of organ donation, but you could easily prepare a lesson focusing on care, heart disease or dementia, as these topics are discussed in *Grey's Anatomy*, too. Word Sketches on *care* and *heart* can be found in Appendix 4.

Furthermore, although the example lesson proposed in this chapter is designed for nursing school students, the tools described can be used for all kinds of professions in vocational schools. When teaching students who want to become mechanics, you could use *Titus* or *Transporter* to build a corpus of language material from that subject area. Similarly, for culinary students, *Chef's Table* may be a useful TV series to create such corpus-based lessons.

5 Caveats and limitations

Sketch Engine offers the possibility to create a 30-day free trial account. This limits teachers' use of the interface to just 30 days. However, within a month, Sketch Engine can be used to create many corpus-based tasks and activities.

Another limitation is that Sketch Engine sets a default word limit to 1,000,000 words for a custom-built corpus. This may seem like a lot at first glance but depending on the focus of the lesson(s), more words might be desirable.

6 Conclusion

This chapter is designed to be a teachers' guide to preparing lessons on staff-patient-communication when teaching English to nursing students. Interactions in hospital settings are explored with a linguistic corpus from the series *Grey's Anatomy* in order to develop teaching material for vocational schools. The episodes consist of dialogues only, thus they are an excellent source of language for students to improve their communications skills. Moreover, the suggested lesson serves as a general example: the range of TV series available offers a wealth of language data that can serve as the basis for creating language teaching materials on a wide array of topics and communicative situations. There are no limits to the teacher's creativity. Although the process of building your own corpus is somewhat time-consuming, task design is in fact easier and faster with the help of a corpus. What's more, corpus-based tasks are based on authentic, native-like language. This is likely to be particularly helpful to non-native teachers and English teachers at vocational schools, who are teaching students learning a profession they are not familiar with themselves. To conclude, both students and teachers can profit from working with a linguistic corpus.

7 Resources and references

Ammende, Rainer, Frank Arens, Ingrid Darmann-Finck, Roswitha Ertl-Schmuck, Brigitte von Germeten-Ortmann, Gertrud Hundenborn, Barbara Knigge-Demal, Uwe Machleit, Christine Maier, Sabine Muths & Anja Walter. 2019. Rahmenpläne der Fachkommission nach § 53 PflBG. https://www.bibb.de/dokumente/pdf/geschst_pflgb_rahmenplaene-der-fachkommission.pdf (26 January, 2020).

Encyclopaedia Britannica. 2018. Grey's Anatomy. <https://www.britannica.com/topic/Greys-Anatomy> (8 February, 2020).

Forever Dreaming. 2020. Grey's Anatomy transcripts. <https://transcripts.foreverdreaming.org/viewforum.php?f=11> (16 June, 2020).

International Council of Nurses (ICN). 2002. Nursing definitions. <https://www.icn.ch/nursing-policy/nursing-definitions> (26 January, 2020).

Niedersächsische Landesschulbehörde. 2019. Grundlegendes zur neuen Ausbildung nach PflBG. <https://www.landesschulbehoerde-niedersachsen.de/themen/berufe-im-gesundheitswesen/pflegefachfrau-und-pflegefachmann/pflegefachfrau-und-pflegefachmann> (26 January, 2020).

Sketch Engine. 2020. What can Sketch Engine do? <https://www.sketchengine.eu/what-can-sketch-engine-do> (26 January, 2020).

8 Appendices

Appendix 1

Vocabulary list

vital

her vital organs

reconstructed

reconstructed organ

removable

removable organs

weakened

weakened organs

abdominal

his abdominal organs

damaged

damaged organ

transplant

transplant organ

internal

internal organs

donor

organ donor

donation

organ donation

recovery

organ recovery

failure

organ failure

fail

organs are failing

bleed

organs are bleeding

donate

donate the organs

reject

reject the new organ

Appendix 2

Cloze exercise

1. Don't make a move until we get our hands on every _____ around the tree trunk.
2. Chief, I know the hypothermia is protective to her _____ but how much longer do you think we can go on like this?
3. This guy could have pretty good life with a _____ that my boy here and I could make tomorrow.
4. We were on a plane, and we just picked up a _____ from a kid who died.
5. He was born with his _____ outside of his body.
6. Grow it in the lab, remove the _____. Replace it with a newer model.
7. That is that the surgery was too intense for his _____ to support.
8. All, all it means is that ah some of the baby's _____ are growing in a sac outside its body
9. Ah, we have an _____ coming in this afternoon from Wilkeson General.

10. He is an excellent candidate for _____.
11. I'm afraid you're gonna go into multi-system _____ if we don't operate.
12. To a critically ill patient, an _____ means a new start, a second chance.

Appendix 3

A conversation about organ donation

Addison: I know Kendra meant the world to you and I know the terrible grief you must be going through right now, but from a medical standpoint, this is a very bad idea.

Mr. Thomas: Well, if we'd wanted your opinion, miss, we'd ask for it.

Mrs. Thomas: It's just...we love our daughter. Please try to understand.

Addison: I do understand. I do. You think if you love her enough or love her baby enough that it will keep her close to you. But she's already gone and if you go forward with this, you're risking...I mean, if Kendra, were to contract an infection, she could pass it on to the fetus and it could cause brain damage...

Mr. Thomas: We're taking our chances.

Mrs. Thomas: Now...they keep organ donors alive after they're...gone. Why not Kendra?

Addison: Organ donors aren't kept alive for six months. Your daughter is brain dead. She can't regulate body temperature, hormone output. These things are very important for a fetus.

Mr. Thomas: You guys can regulate that stuff, right?

Addison: We could try. I...

Mr. Thomas: Well then try, damn it! What the hell's the problem here?

Addison: The problem, as I see it, is you want to use your daughter's corpse as an incubator. That's the problem.

Derek: Ok, I'm sorry. Would you excuse us?

Source: <https://transcripts.foreverdreaming.org/viewtopic.php?f=11&t=6770> (16 June 2020)

Appendix 4

Word sketches of *care* and *heart* can be generated and worked with as described in Section 4.1.2.1.



Fig. 15: Word sketch of the word care generated with the 'Word Sketch' function.

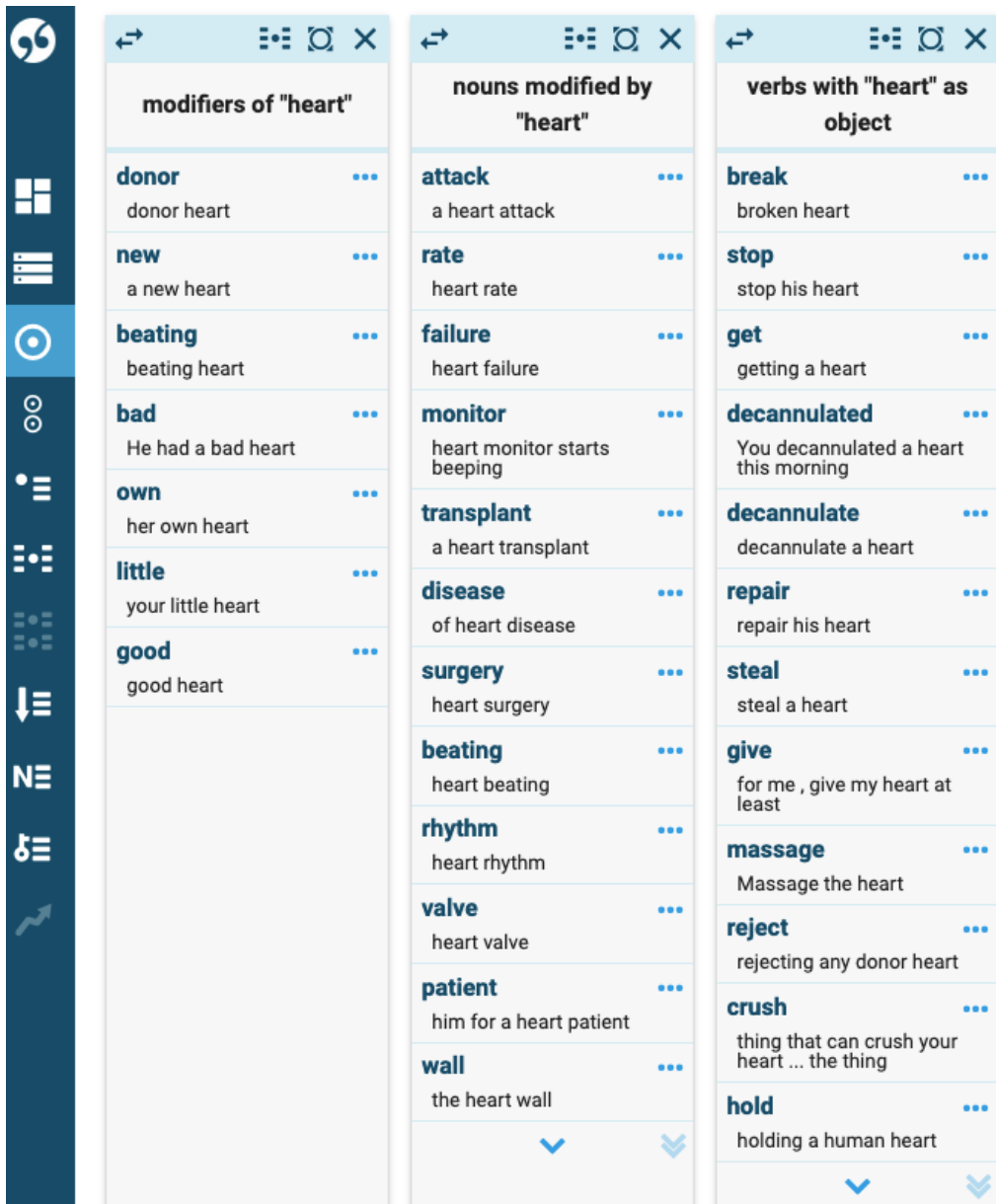


Fig. 16: Word sketch of the word heart generated with the 'Word Sketch' function.

16. Creating teaching materials for nursing schools using medical TV series

KATINA VORBRINK AND ANNA FANKHAUSER

Unlike the other chapters of this e-book, which were created as part of three iterations of a project seminar (see “About the project”), the present chapter was developed on the basis of **Katina Vorbrink’s M.Ed. thesis** (Osnabrück University, 2020), which was supervised by Alexander Bergs and Anna Fankhauser. Since this thesis project fits in this OER project very well, with Katina’s approval, it was adapted by Anna for inclusion in this e-book, in the hope that the materials created as part of Katina’s thesis may be used and adapted by English teachers at nursing schools in Germany and beyond.

Abstract

In response to the growing shortage of skilled workers in the nursing sector, nursing education in Germany has recently undergone major changes. As part of a large-scale education reform that came into force at the beginning of 2020, the former tripartite system of nursing apprenticeships (adult nursing, children nursing, and geriatric nursing) were merged into one apprenticeship profile. Restructuring the nursing curriculum involved the implementation of general subjects such as English, thus leading to the need for new adequate English language teaching materials. This chapter explores how corpus linguistics can serve as a helpful tool for the development of teaching materials in the specialized field of nursing education.

The promotion of students’ communicative skills is the main focus of EFL teaching in the nursing education context. Hence, a spoken corpus for the medical context (Social English for Medical Context - SEMC) was created based on transcripts from the three medical television shows *Emergency Room*, *Dr. House*, and *Grey’s Anatomy*. The selected shows proved to be suitable with regard to medical accuracy and their potential to promote communicative skills. Five corpus-informed sets of English as a Foreign Language (EFL) teaching materials for nursing schools were developed as a first step towards establishing English as a general subject of nursing education in Germany.

1. Background

In the course of the recent major reform of nursing education in Germany and the subsequent implementation of English as a general subject in the nursing curriculum, the need for new, adequate English language teaching materials for this particular vocational school context arose. Since vocational education aims to meet the requirements of the learners’ future professional practice (Belcher 2006: 134), teaching materials are preferably based on naturally occurring language produced in authentic situations of the respective work environment. The present chapter explores how corpus linguistics can serve as a helpful tool for the development of authentic teaching materials in the specialized field of nursing education

(for a detailed account of the benefits of corpus linguistics in language teaching in general and in the vocational school context in particular, see chapters “About the project”, “Developing ESP learning activities for vocational education with a custom virtual corpus” and “Exploring hospital interactions with Grey’s Anatomy”).

In order to identify the specific requirements for English teaching materials in the nursing education context, the education reform and its implications for the educational practice, as well as formal prerequisites, learners’ language needs, characteristics of potential learning groups and learning targets will be analyzed in the following.

1.1 Nursing care reform and educational practice

The nursing profession in Germany is a relatively young profession in comparison to other countries such as the United States (Menche 2007: 24). Education in nursing is constantly changing and evolving. The most recent development is the nursing care reform of 2017, which came into effect on the first of January 2020 (Jürgensen 2019: 7) as a response to the demographic change on the one hand and the lack of nursing staff on the other (Hamar/Kuckeland/Schneider 2019: 152). The legislative reform merges the former tripartite system consisting of three separate nursing training programs (one for each patient age group: geriatric, pediatric, and adult nursing) into a single job profile referred to as the generalized nursing apprenticeship (Jürgensen 2019: 7). Duration of the apprenticeship amounts to three years in total and, in contrast to other countries, is subject to a particular dual training system that combines theoretical components (taught at vocational schools) with practical training on the job in different clinics and outpatient clinics (Pflegerberufereformgesetz – PflBRefG 2017: 2584). More specifically, the three-year program is divided into specialized classes, general subjects, and practical training. One concrete objective of the education reform is to provide trained nurses with a qualification that enables them to work in other European countries (Bundesministerium für Gesundheit – BMG 2018: n.p.).

The reform implies reformulating the nursing curriculum with regard to the technical subject contents on the basis of the framework curriculum of the state school authority. Each nursing school is required to transfer these guidelines into its own curriculum. In addition to the restructuring of the technical subject contents, the implementation of four general subjects, English, German, Politics, and Religious education was also an essential aspect of the reform.

The guidelines of the local Ministry of Education (Niedersächsisches Kultusministerium – NKM 2009a) serve as a common ground for English teaching in nursing education. The guidelines for English refer to the subject as ‘English / Communication’, which already indicates a particular focus on promoting the students’ communicative competence.

1.2 Formal prerequisites for teaching materials and learners’ language needs

As briefly mentioned above, the Ministry of Education’s guidelines for English in the context of nursing education refer to the subject as ‘English / Communication’, which stresses the focus of the communicative competence in vocational schooling. This specific feature of English in nursing education, as well as the characteristics of potential learning groups and specific learning targets have to be taken into careful consideration when approaching the creation of suitable EFL teaching materials.

English is of great importance for nursing trainees for several reasons. Miyake/Tremarco (2005: 32) draw attention to the fact that, due to its status as an international *lingua franca*, English in particular can serve as a linguistic bridge in communication with patients and their family members. This is essential

because the number of patients hospitalized in countries with official languages other than their native languages is rising worldwide. It is important for nurses to be able to successfully communicate in order to ensure comprehensive care. Listening and speaking skills have to be carefully addressed to enable nurses to engage in professional communication (Akbari 2016: 24). Klitzing et al. (2007: 42) add that the majority of medical studies are published in English. Receptive skills in general should therefore also be focused on in English language teaching for nurses. However, this aspect of language learning is not addressed in this chapter.

The English curriculum for nursing schools is based on two important sources: the framework directive of the Ministry of Education in Lower Saxony (NKM 2009a) as well as existing research evidence regarding learners' language needs (Guhde 2003; Miyake/Tremarco 2005). It is important to note that the ministry's guidelines are formulated in a fairly general manner and that it is up to every school to create a school-specific curriculum for the subject.

The guidelines formulated in the framework directive of the Ministry of Education in Lower Saxony consist of four central competence areas, namely 'Reception', 'Production', 'Interaction', and 'Mediation'. The guidelines describe each competence in detail and list examples of possible teaching materials and methods that may be applied in class. The competence 'Reception' is defined as the students' ability to understand common texts specific to the vocational context (NKM 2009a: n.p.). The Ministry of Education states that this skill describes the idea that students are able to follow and understand basic lines of discussions and speeches. The list of example texts includes technical texts and newspaper articles. Texts to address listening skills include communicative situations, podcasts, and video sequences. 'Production' describes the students' ability to write texts and statements in the target language depending on the situation and purpose. These can include answers to questions of text comprehension or filling out forms. Spoken communication is central to this competence and refers to giving a presentation or to reporting something. The third competence is 'Interaction'. The individual learner should be able to ask for relevant information in connection with typical professional tasks. The students should be able to apply simple communication techniques. Taking a patient's medical history is a good example of this competence. The final competence described in the framework is 'Mediation'. It does not only refer to the translation of language material into the target language but also to the ability to orally summarize texts and utterances in English (ibid.).

The emphasis on particular skills varies according to the year of schooling. Following the recommendations of Miyake/Tremarco (2005: 33), in the beginning the focus lies on 'Reception' and 'Production', skills closely related to the teaching of technical terms and vocabulary acquisition. In the course of the apprenticeship, the focus shifts towards the competence of 'Interaction', emphasizing the students' communication skills more strongly.

In addition to the language needs of the nursing students it is also important to pay attention to the rising need for media skills. Trübswetter/Figueiredo (2019: 345) looked at the potential of computer-based learning methods in nursing schools. They underline the need for schools to foster students' technical competence in the new nursing curriculum in order to remove barriers and to build up confidence in using electronics. This demand results from changes in the work environment and increasing technology use that has to be mastered by nurses in the clinical setting, e.g., the electronic patient file (Trübswetter/Figueiredo 2019: 348).

In summary, it can be said that the competences required for nursing students are centered on social interaction and communication in the clinical setting. According to Miyake/Tremarco (2005: 23), the language type needed to facilitate communication is the so-called 'Social English' as opposed to 'Technical English', which is considered less important in nursing education. The term 'Social English' refers to language likely to be used in interactions with different individuals, whereas 'Technical English' is characterized by vocabulary likely to be encountered in the clinical setting (Miyake/Tremarco 2005: 24). The authors (Miyake/Tremarco 2005: 33) suggest limiting 'Technical English' in the classroom of nursing education to basic technical vocabulary, such as anatomical terms and clinical symptoms. 'Technical

English' can build on learning in previous years of schooling (Miyake/Tremarco 2005: 26). They conclude by stating that 'Social English' becomes increasingly important in the course of the apprenticeship as it enables communication with patients and other individuals in the clinical setting. Miyake/Tremarco (2005) strongly suggest focusing on 'Social English' in the last year of the nursing education program. A closer look at the nursing curriculum for the subject 'English / Communication' of a nursing school in Lower Saxony (see table 1) shows that these recommendations have been followed. Towards the last year of the program, the emphasis shifts from the receptive competence and 'Technical English' to the competence of 'Interaction' centered on 'Social English'.

Table 1: Overview of the English curriculum of a nursing school in Lower Saxony (Vorbrink 2020: 29)

Curriculum English / Communication (90)				
Competence area	Number of lessons	1 st school year (30)	2 nd school year (28)	3 rd school year (32)
Reception	26/90	14	8	4
Production	30/90	10	10	10
Interaction	24/90	4	6	14
Mediation	10/90	2	4	4

Although the guidelines of the Ministry of Education (NKM 2009a: n.p.) describe important competences of the nursing students, they are very vague and lack detail. Therefore, additional sources had to be consulted in order to determine precise language needs for nurses.

Cameron (1998) looked at language needs for nursing students in class and clinic. His study is recurrently used to serve as a basis for nursing curricula (Bosher 2013: 264; Cameron 1998: 215). Cameron (1998: 205) based his needs assessment on a needs-definition that is understood in four ways: analytic, democratic, diagnostic, and discrepancy. A reference group, for example nursing educators and students, forms the basis of the analytic and democratic needs perspectives. By consulting a reference group, one is more likely to identify useful needs for the learners' group and to minimize the risk of overlooking essential needs. For example, educators and students articulated the need for extra practice in class (Cameron 1998: 206). The discrepancy approach of Cameron's needs definition refers to observed differences between the actual performance of a student and the required action (ibid.). The diagnostic needs definition aims to rank the identified needs with regard to frequency and relevance of the respective need (Cameron 1998: 207). Clearly, Cameron's underlying needs definition is comprehensive and takes several aspects of language learning into consideration.

Cameron (1998: 215) identifies five major needs of nursing students, which are displayed in Fig. 1. As mentioned above, the determined needs 'serve as suggestions for course content' (Cameron 1998: 203) and are recurrently cited by other authors (Bosher 2013: 264; Nation 2013). The first need that Cameron (1998: 208) defines is 'Speech Production Accuracy'. It is subdivided into the four categories pronunciation, vocabulary, grammar, and discourse. 'Academic Performance', as the second need, consists of five individual skills, such as writing and listening. The third need, 'Clinical Performance', is split into several smaller parts, of which three are of special interest to this chapter: getting information, transmitting information to a patient or a family member, and social interaction with different individuals at work. 'Dialect (Cultural) Variation', is the fourth need that Cameron (1998: 214) describes. It relates to cross-cultural awareness of nurses with regard to a multicultural patient group and colleagues. And finally, the need 'Inferencing Skills' is based on Cameron's definition of inference, which refers to observations that may be drawn with respect to a speaker's communicative intention (ibid.).

Not all of the needs identified by Cameron will be addressed in this chapter. The designed teaching materials in this chapter are exclusively developed to satisfy the needs one and three. Fig. 2 gives an overview of the two selected need statements and intends to provide a better understanding of each need.

I. Speech Production Accuracy



II. Academic Performance



III. Clinical Performance



IV. Dialect (Cultural) Variation



V. Inferencing Skills: Language-Based and Social-Based



Fig. 1: Overview of the needs statements for nursing students in class and clinic (Vorbrink 2020, based on Cameron 1998: 215)

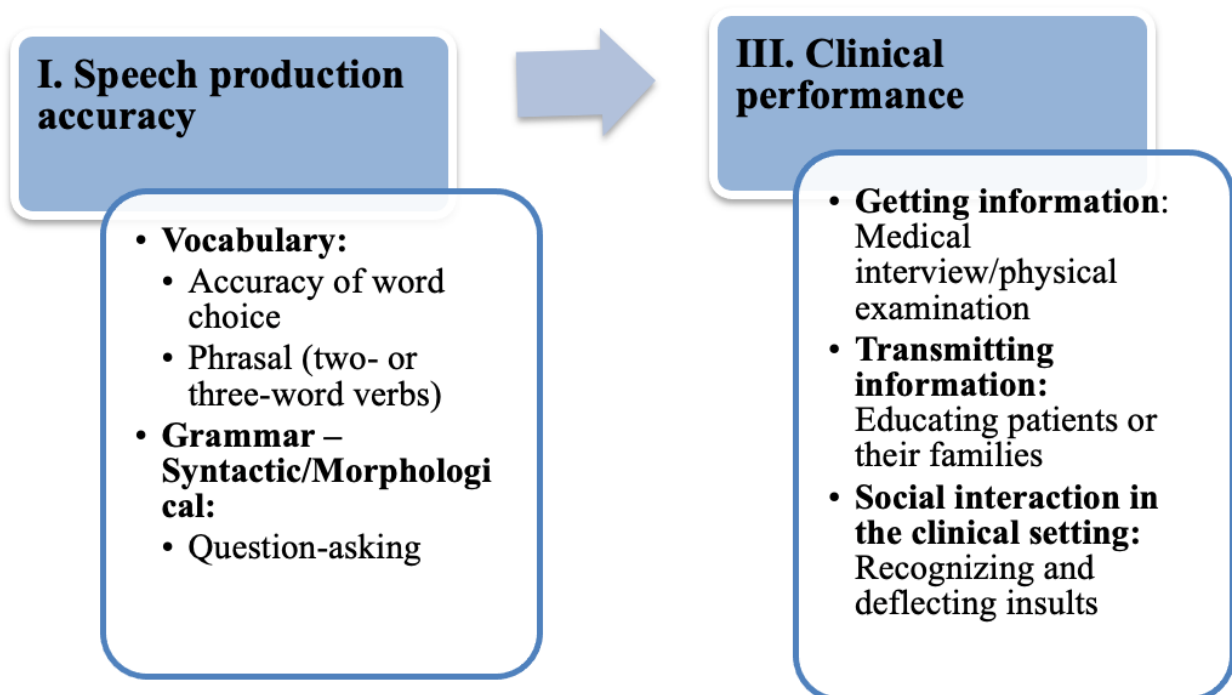


Fig. 2: Selected needs for the teaching set (Vorbrink 2020, based on Cameron 1998: 208-214)

A further source consulted to identify specific language tasks with which nursing students will be confronted at work are the results of a study conducted by Epp/Lewis (2009). The authors observed nurses in the clinical setting for a certain period of time and identified a range of essential and recurring language tasks (Epp/Lewis 2009: 308). The majority of communication takes place with patients, also referred to as clients (see Fig. 3). Individual language tasks performed by nurses range from 'giving instructions' over 'asking for information' to 'making small talk'. All identified language tasks with the associated percentages of the time spent on the tasks are shown in Fig. 4.

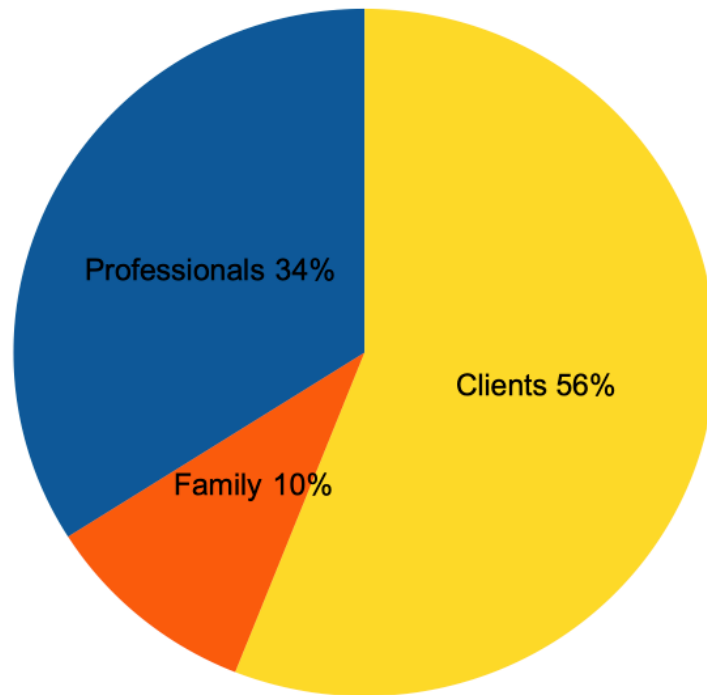


Fig. 3: Situational use of language in the clinical setting by nurses (chart based on Epp/Lewis 2009: 308)

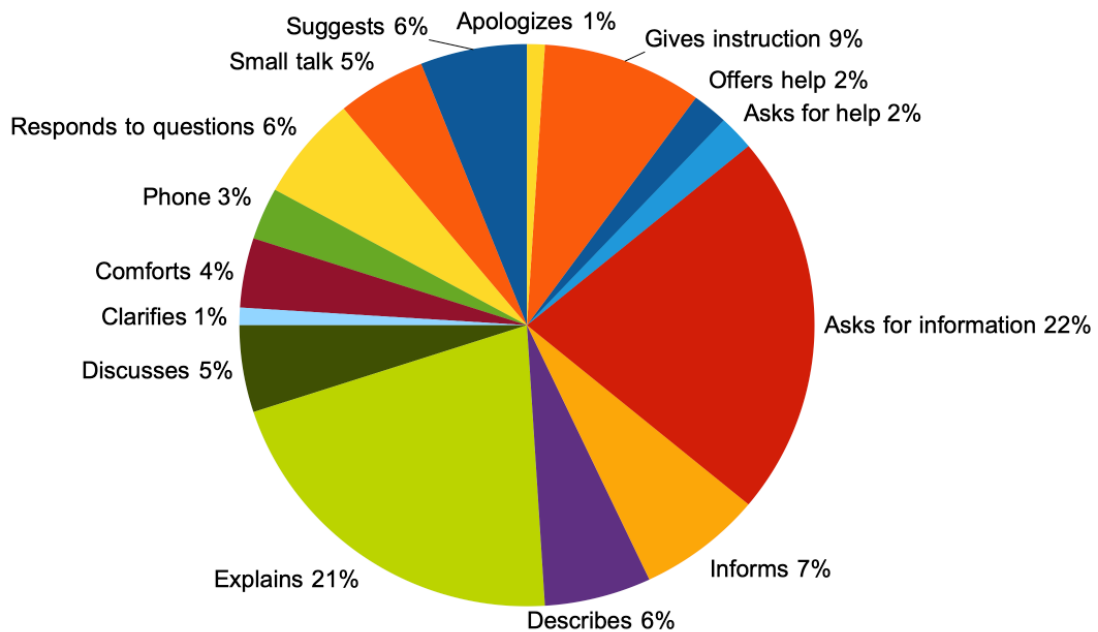


Fig. 4: Types of nursing language tasks and the percentage time spent on each task (chart based on Epp/Lewis 2009: 307)

1.3 Characteristics of potential learning groups

The requirement for a nursing apprenticeship in Germany is the middle school qualification known as the 'Realschulabschluss' (NKM 2009b: n.p.). The guidelines of the Ministry of Education (NKM 2009a: n.p.) define the B1 level as the language proficiency level of students who have completed secondary education. According to the ministry's guidelines (ibid.), the students have acquired a general language vocabulary and grammar rules in their previous ten years of general schooling. Experience shows that each nursing class regularly contains a small number of high school graduates who acquired a higher level of language proficiency during their thirteen years of schooling (B2/C1) (Niedersächsischer Bildungsserver – NibiS 2020: n.p.). For an educator, it is crucial to keep this heterogeneity in mind in terms of internal differentiation. Generally, this term refers to the idea that teaching content and materials are carefully tailored to the different student groups with regard to their level of proficiency. There are different teaching methods that can be used in combination with each other in order to support the different learning needs of a group. Donath/Klemm (2013: 145), for example, name the method *peer correction* and suggest letting students work in small groups to create a safe learning environment.

With the new care reform, it is possible for students who have finished a two-year apprenticeship as nursing assistants to skip the first year of the generalized nursing apprenticeship and to enter the second year of schooling (Ministerkonferenz – MK 2020: 3). This reduction of the number of training years is based on the idea that these students have attained a solid foundation of nursing and work experience in advance (Jürgensen 2019: 14). The school curriculum for nursing assistants also includes a moderate level of English language teaching (Jürgensen 2019: 62). Therefore, it is reasonable to assume that these students also have a satisfactory general language vocabulary and already acquired a moderate knowledge of technical vocabulary.

Generally, nursing students show strong motivation for the apprenticeship because it is a first step in their professional career. However, in recent years German nursing students' motivation and independent learning skills seem to have decreased (Baumgarten/Ayerle 2016: 56). The authors (Baumgarten/Ayerle 2016: 53) point out that nursing students face increasing demands that arise from demographic change and growing complexity of work processes. The students seem distracted by new media more than ever (Baumgarten/Ayerle 2016: 56). Trübswetter/Figueiredo (2019: 345) agree and go on to predict that new media will become a component of nursing education for years to come. Incorporating media competence into the curriculum is a crucial measure to promote a reasonable use of technology in the future. They go on to underline that these actions represent a necessary reaction to rapidly changing work processes in healthcare (ibid.).

To sum up, the creation of corpus-based teaching materials for nursing schools in Lower Saxony is shaped by the binding guidelines of the Ministry of Education (NKM 2009a: n.p.) for the subject 'English / Communication', nurses' language needs as identified by Cameron (1998: 215), frequent language tasks as identified by Epp/Lewis (2009: 308), as well as by the individual characteristics of learners in the nursing context.

2. Methodology and procedure

The previous section has shown that social interaction and communicative skills are the focus of English teaching in the context of nursing education. The corpus-based teaching materials designed for teaching English to future nurses in the course of the present study are therefore centered on 'Social English'. Existing medical corpora have failed to sufficiently cover this particular language type. Consequently, a corpus specifically representing 'Social English' had to be compiled.

In general, corpora that have been derived from medical contexts are of particular interest for the material development for future nurses. Sketch Engine, a web-based tool to create and investigate preloaded and individually compiled language corpora, offers a large corpus in this field, the Medical Web Corpus (MWC). The MWC contains over 33 million words predominantly from medical websites. A wordlist analysis of nouns has shown that this corpus contains a large number of broader technical terms in health care and medicine, such as *infection*, *therapy*, and *diagnosis*. It also contains more specific words from the medical field with lower frequencies, such as *hemorrhage*, *anemia*, and *dilatation*. The MWC is well suited to generate vocabulary lists for the nursing students. Nation (2013: 305) stresses that wordlists are a possible tool for teachers to help the students with vocabulary acquisition. However, a closer look at the MWC makes clear that this particular corpus, with its large variety of technical vocabulary, is not suitable for promoting the nursing students' communicative skills during their apprenticeship. Therefore, a smaller specialized corpus was compiled as suggested by Aston (2001: 37), who points out that from a pedagogical perspective, it can be beneficial to turn to a small specialized corpus instead of a large general one.

To allow for the promotion of the learners' communicative skills, a corpus for social English entitled 'Social English for Medical Context (SEMC)' was compiled using the online corpus tool Sketch Engine. The SEMC contains more than 920.000 words taken from 225 television transcripts and can be considered an addition to the MWC.

The methodological approach to generate a spoken corpus from television shows is based on the findings of Levshina (2015: 23), who points out the similarities between online subtitles and daily spontaneous conversations in American and British settings. Apart from the authenticity of the linguistic interaction, the technical correctness of the selected medical television shows is of importance here. A review conducted by Hoffman et al. (2018: 202) emphasizes the fact that physicians participate in the writing process of the scripts and that this ensures the correctness of the medical conditions and techniques in the shows. The authors go on to suggest that teachers may use television scripts for educational purposes in class (Hoffman et al. 2018: 215).

The next step was to determine suitable medical television shows for the creation of the subcorpus. The final selection of fictional medical shows is based on available studies on this particular topic. Hoffman et al. (2018: 202) examined *Emergency Room (ER)*, *Grey's Anatomy* and *Dr. House*. The realistic medical representation in *ER* is particularly highlighted in various studies (Oswald 2009: 112). Krüger-Brand (2006: 2350) goes on to report that *ER* contains many scenes with typical everyday hospital life issues, such as tensions between patients and their family members or communication with patients in general. Wong et al. (2009: 2) draw attention to the effective communication skills presented in scripts from *Dr. House* and *Grey's Anatomy*. Leucht (2013: 81) confirms the medical accuracy in *Dr. House* and *ER*. Nevertheless, it is important to point out that the selected medical television shows predominantly focus on the medical staff rather than on the nursing staff. Therefore, the television scripts used for the SEMC may cause a shift of emphasis in the communication presented.

The television scripts can be obtained from specific websites. The transcripts of *Grey's Anatomy* were taken from Forever Dreaming Transcripts (2020), the transcripts of *Dr. House* from LIVEJOURNAL (2020) and those of *ER* from the website Simply Scripts (2006). The transcripts were selected randomly by choosing every other transcript of the corresponding website with exception of *ER*. For this show, a total of six transcripts was available and therefore all six were included. In total, 175 transcripts of *Grey's Anatomy* and 44 of *Dr. House* were selected.

3. Corpus-based teaching materials for the subject ‘English / Communication’

The following section describes five sets of teaching materials for nursing schools in Lower Saxony for the general subject ‘English / Communication’. The time required for each set is a 90-minute classroom session. The description of each set is prefaced by a brief identification of the learners’ needs and the linguistic relevance of the corpus-based worksheet. Furthermore, each set is embedded in the nursing curriculum with close connection to the underlying framework in the curriculum of the Ministry of Education (NKM 2009a: n.p.). Each subsection closes with ideas for the further use of the teaching materials in following classroom sessions. The materials cannot be seen in isolation but rather in a succession of consecutive classroom sessions. One session can serve to introduce a topic and a subsequent session either to explore the topic in more detail or to provide additional time for practice (Hutchinson/Waters 2010: 108-109).

Table 2: Overview of corpus-based teaching materials for the general subject ‘English / Communication’ (Vorbrink 2020: 37)

Classroom session	School year	Competence area according to NKM (2009a)	Learners’ need according to Cameron (1998)
1 Giving instructions vs. informing someone	1	<ul style="list-style-type: none"> • Reception • Production 	I. <u>Speech production accuracy</u> – Vocabulary <ul style="list-style-type: none"> • Phrasal
2 Offering and asking for help	2	<ul style="list-style-type: none"> • Reception • Interaction 	I. <u>Speech production</u> – Grammar – syntactic <ul style="list-style-type: none"> • Question-asking
3 Belly vs. stomach	2	<ul style="list-style-type: none"> • Reception • Production 	I. <u>Speech production accuracy</u> – Vocabulary <ul style="list-style-type: none"> • Appropriateness of word choice
4 Getting information – Patient interview	3	<ul style="list-style-type: none"> • Interaction • Reception 	III. <u>Clinical performance</u> – Getting information <ul style="list-style-type: none"> • Patient interview <ul style="list-style-type: none"> ◦ Medical history
5 Additional contexts of <i>help</i> (v/n)	3	<ul style="list-style-type: none"> • Interaction • Production 	I. <u>Speech production accuracy</u> – Vocabulary III. <u>Clinical performance</u> <ul style="list-style-type: none"> • Social interaction with different individuals in the clinical setting

3.1 Giving instructions vs. informing someone

The first worksheet targets vocabulary acquisition in the nursing students' first year of schooling. With reference to the guidelines of the Ministry of Education (NKM 2009a: n.p.), this worksheet can be used to facilitate the competence areas 'Reception' and 'Production'. As pointed out above, two of the core language tasks of nurses are 'giving instructions' and 'informing a patient' (Epp/Lewis 2009: 308). In order to form correct sentences in this regard, nursing students have to adequately use typical phrases for these tasks (Cameron 1998: 209). The worksheet therefore aims to teach relevant patterns that are associated with 'informing someone' and 'giving instructions' in the clinical setting. These patterns, e.g. 'I need you to' or 'I have to', first have to be learned as idiomatic chunks (Aston 2001: 15) before they can be fluently applied in varying communicative situations. This exercise may not only help the nursing students to acquire relevant phrases, but also helps to raise learners' awareness of the particular context these phrases are used in (Aston 2001: 16).

The worksheet (Fig. 1 + II) is divided into two parts. Firstly, common phrases for 'giving instructions' and 'informing someone' are introduced with the help of concordance lines from the SEMC. Sketch Engine was used to generate a 2- to 6-gram list on the basis of which phrases commonly used to perform recurrent language tasks were identified. The n-grams which ranked highest (count 281), e.g. 'I have to', were mixed with n-grams of mid-frequency (count 92), e.g. 'I want you to'. An additional collocation search provided the concordance lines for the first exercise of the worksheet. The relevant phrases appear in bold to emphasize the targeted patterns in the concordance lines. Nation (2013: 147) suggests the classification of items into groups as a possible activity for students to engage with concordance lines. Following Nation's suggestion, the students are asked to read the concordances and to categorize the phrases in bold into two groups, the phrases that can be used to give information and the ones to give instructions.

English / Communication	Verbal Phrases	
Date:	Exercise Sheet	

1. As a nurse, one of your key roles is to give information and instructions to patients and their family members. Read the concordance lines below and group the phrases in bold into two categories: Which ones can you use **to inform somebody about something** and which ones can you use **to give instructions**?

1. do not move your head. I need you to keep perfectly still. Don't move.
2. Christina: I'm sorry, but I have to ask one more time.
3. clean up with this patient. I have to take you up to surgery now, and I will ...
4. down beside an ambulance. I want you to slowly open your eyes and try to stay calm, ...
5. I'm sorry. I just ... wanted to make sure that you're okay. Are you okay?
6. Foreman: I'm sorry, but I'm gonna have to ask you to sign this.
7. Who are you with? Ok, I want you to stay right here, ok?
8. Ok sir, you're sweating. I need you to sit down, get back in the bed.
9. You wait right there! I want to talk to you!
10. Mm. (softly) I want you to continue to move your fingers while in the cast.
11. Cuddy: I need you to lie on your side. And hold your knees.
12. Okay. I want to ask you one more time... Are you absolutely sure?
13. could be a lengthy process. We're gonna have to run more tests.
14. Please stop! Erica: Mr. Arnold, I need you to calm down so that your heart can calm down.
15 They're not. Excuse me. I have to go check on my patient.

Phrases you use to inform a patient or a family member about something	Phrases to give a patient instructions

Fig. 1: Teaching set number 1 "Phrases for giving information and instructions" – Part 1 (Vorbrink 2020: 67; concordance lines: SEMC)

Secondly, the students are asked to practice the use of the phrases by completing sentences on the exercise sheet. This exercise allows the students to recall the items when the context is presented (ibid.). The sentences chosen for exercise 3 of the worksheet were compiled from various concordance lines of the same collocation search in the SEMC corpus. The concordance lines present the target phrases in authentic and diverse contexts. As the corpus is compiled from TV transcripts and predominantly represents communicative interaction, some concordance lines contain informal expressions such as 'gonna have to'. These underline the language authenticity but need to be addressed in class to increase students' awareness of the formal vs. informal tone of certain words and phrases (see exercise 2 on the worksheet).

2. “Gonna have to”

You have probably noticed the expression “gonna have to” in the lines above. Pair up with your neighbor and talk about the following questions:

- What would you say about its tone? Is it rather formal or informal?
- In which kind of conversation is it appropriate to use this form?
- What is the formal version of the expression?

3. Exercise: Giving Instructions and Informing Someone

Complete the sentences by filling in the blanks with the correct phrases. Are you informing the patient about something or are you giving precise instructions? The first sentence is given as an example.

1. Nurse: I am here to help you. I need you to turn around, so I can check your back for a skin rash.
2. Patient [coughs]: I have a high temperature and a sore throat. Nurse: _____ cough for me, okay? And _____ take your temperature.
3. Nurse: It's an infusion. _____ give it to you. All the rest of your symptoms should go away.
4. Patient [moans in pain]. Nurse: Look at me. _____ keep still while I dress the wound.
5. Nurse [holds out a cup with pills]: _____ to take them. They should help to control the pain.
6. Patient [tries to sit upright in bed]: I need to go to the toilet. Nurse: Sir, wait. I'm going to help you. _____ lie on your left side and then I can get you out of bed over its edge.
7. [Patient bell is ringing] Nurse: I'm sorry. I have to go now. _____ check on another patient.
8. Nurse: Good morning. I just _____ you were all right in here. Did your symptoms go away?
9. Patient: I am allergic to certain pain medications. Nurse: Ok, I am _____ get Dr. Amber. He will ask you more questions about your medical history.
10. Nurse: _____ talk to you about your seizures. Did you ever have problems like this before?

Fig. II Teaching set number 1 “Phrases for giving information and instructions” – Part II (Vorbrink 2020: 68; concordance lines: SEMC)

The teaching set can be continued in following class sessions by using the acquired phrases in the form of a practical application in the nursing school's skills. The students get together in pairs and improvise a patient-nurse-interaction by using the acquired phrases for 'giving instructions' and 'informing a patient'. This sequence allows the students to practice and strengthen the acquired verbal phrases in connection with operational nursing tasks. The previous class session functions as a language learning task and the following session as an activity-oriented communicative task (Legutke 2013: 94). The combination of subject teaching and language teaching facilitates the development of the learners' vocational competence (Legutke 2013: 116; NKM 2009a: n.p.).

3.2 Offering and asking for help

This teaching set is based on the language tasks of 'offering help' and 'asking for help' as identified by Epp/Lewis (2009: 308). The targeted competence areas from the guidelines (NKM 2009a: n.p.) are 'Production' and 'Interaction'. Typically, nurses perform these tasks in contact with patients but also with family members and other professionals in the work setting (Epp/Lewis 2009: 307). The aim of the exercises of this set is to improve the 'Speech production accuracy/appropriateness' of the nursing students with regard to their grammar structure of asking questions (Cameron 1998: 209). The focus is not on a specific type of question. Instead, it aims at making the students familiar with different ways of asking for and offering help because actual interaction with patients and colleagues does not exactly correspond to the arranged dialogues presented in many conventional schoolbooks (O'Keefe/McCarthy/Carter 2007: 22; Gabrielatos 2005: n.p.). The lack of authenticity in textbook conversations has been linked to poor communication skills at all proficiency levels (Savova 2018: 85). To address this problem, concordance lines have been selected from the SEMC corpus and used as a foundation for the exercise. The broad context with rich exposure to the use of *help* (n/v) targets the communicative competence of the students (Belcher 2006: 148).

In order to extract suitable examples, a concordance search with the lemma *help* (n/v) was performed in Sketch Engine and revealed a total of 887 concordances. By skimming the concordances for suitable contexts, e.g., significantly asking for and offering help, the number of concordances was minimized to a manageable number of 40 for the students to analyze in class (Fig. IV + V). The aim of the worksheet is based on Gabrielatos (2005: n.p.). Students are asked to identify different ways of using *help*. In order to expose the students to authentic interactions, the chosen concordance lines display different contexts for the use of *help*. The concordance lines include instances of 'asking for help', 'offering help', and also occurrences of 'calling for help'. The learners should be able to notice the differences. In addition, the students are asked to think of reasons why someone would ask for help in different ways. Students are expected to understand that a call for help is usually linked to an emergency situation and requires a rapid response on their part (Menche 2007: 15). The chosen 'Think!-Pair!-Share!' teaching method is a form of cooperative learning, which initiates the exchange of ideas and facilitates the students' communicative competence by maximizing the speaking time of each student (Grieser-Kindel/Henseler/Möller 2017: 186). This approach meets the recommendations by Donath/Klemm (2013: 145) to react to the possible heterogeneity of the learners.

English / Communication Date:	Help (v/n) Exercise Sheet	
----------------------------------	--	--

1. Think for yourself!

- Read the lines below. Work out the way that *help* is used. Use a dictionary if necessary.
- Decide if *help* is used in the same way in all of the sentences.
- Try to group the ways *help* is used.

2. Pair up!

- Compare with your neighbor. What is different and what is similar?

3. Share!

- Be prepared to share your findings with the group.

4. Practice! – Role play

- Team up with your neighbor and think about a situation on the ward where you have been asked for help or offered help to someone.
- Write a dialogue for that situation. Use the concordance lines as guidance for writing the questions.
- Now, decide on who will play which role and practise your role play a few times. The goal is to memorize your text and to speak without your notes.
- Be prepared to act out your role play in front of the class.

Fig. III: Exercise sheet for teaching set number 2 "Offering and asking for help" (Vorbrink 2020: 69, based on Grieser-Kindel/Henseler/Möller 2017: 186-187)

1	tripod.com	LIZABETH: He's in surgery, is it something I can	help	you with? </s><s> MR. </s><s> GARDENER: W
2	tripod.com	ou expect me to do? </s><s> YOSH: Could you	help	me get him out? </s><s> PETER: Uh huh, I don't
3	tripod.com	aby. </s><s> There's no way. </s><s> Can you	help	me? </s><s> DOUG: Yes I can, I'll be right back.
4	archive.org	ts. </s><s> MALUCCI: Sorry. </s><s> Need my	help	? </s><s> CHEN: I've got it. </s><s> Pan to and
5	archive.org	know it hurts. </s><s> I'm a doctor, I'm going to	help	you, all right? </s><s> Okay. </s><s> Look at m
6	archive.org	ound) Yeah. </s><s> I'm Dr. Kovac, I need your	help	. </s><s> Can you make room for this girl? </s><
7	archive.org	tres blood mass. </s><s> PLUMBER: Need any	help	? </s><s> LUKA: No, I got it. </s><s> Thank you
8	archive.org	's><s> CHEN: All right, um... </s><s> You need	help	back on the bed? </s><s> CARTER: No, I'm good
9	archive.org	ind your lounge area. </s><s> MALUCCI: Can I	help	you? </s><s> We don't usually let people wait in
10	archive.org	STEVE: Well, while you're at it, maybe you can	help	me out and read a few plain files. </s><s> MAR
11	archive.org	the doorframe. </s><s> BENTON: Here, let me	help	. </s><s> CARTER: I got it, I got it. </s><s> He l
12	foreverdreaming.org	s><s> Woman: Help! </s><s> I need immediate	help	. </s><s> Bailey: What's the problem? </s><s> V
13	foreverdreaming.org	ike it easy. </s><s> Is there anything I can do to	help	? </s><s> Bailey: Uh, yeah for starters you can h
14	foreverdreaming.org	ristina (calls from the other OR): Hey, somebody	help	! </s><s> Help here! </s><s> (They all sprint fro
15	foreverdreaming.org	n the other OR): Hey, somebody help! </s> <s>	Help	here! </s><s> (They all sprint from the room and
16	foreverdreaming.org	rgery. </s><s> Cris... </s><s> I ... I need you to	help	me Cristina. </s><s> (She hands Will over to Cri
17	foreverdreaming.org	posted. </s><s> OMAR: Hello? </s><s> I need	help	please. </s><s> I know you can hear me. </s><s>
18	foreverdreaming.org	aylor! </s><s> CRISTINA (shouting): Someone	help	here please! </s><s> INT. </s><s> TAYLOR TR
19	foreverdreaming.org	gger by the second. </s><s> DEREK: You need	help	? </s><s> HAHN: I need four hands and I need e
20	foreverdreaming.org	ommunity) </s><s> George: Do you need some	help	? </s><s> (The man replies in Pennsylvania Du

Fig. IV: Concordance lines for teaching set number 2 "Offering and asking for help" – Part I (SEMC; correction of line 10 from "films" to "files")

21	foreverdreaming.org	ate directions) </s><s> George: You need some	help	? </s><s> Paramedic: Nothing more I can do he
22	foreverdreaming.org	walks by) </s><s> George: Callie, I need some	help	. </s><s> Callie: I can't, I'm busy. </s><s> Georg
23	foreverdreaming.org	><s> Damn it. </s><s> (Loudly) Can I get some	help	over here? </s><s> Sweetie, listen to me. </s><
24	foreverdreaming.org	rtgomery. </s><s> Addison: Yes. </s><s> Can I	help	you with something? </s><s> Jeff: I'm Jeff Pope
25	foreverdreaming.org	'lease, just... </s><s> I can't...help me, help me,	help	me! </s><s> (Meredith literally cuts the wedding
26	foreverdreaming.org	e and Camille) </s><s> Doctor: We need a little	help	here! </s><s> Adele: Oh, thank god. </s><s> Sc
27	foreverdreaming.org	erty SUV squeals into the Bay.) </s><s> Woman:	Help	us! </s><s> Please help! </s><s> Phillip: Help n
28	foreverdreaming.org	toothbrush ... and more knife. </s><s> Can you	help	me? </s><s> You see, I'm gonna get a wish from
29	foreverdreaming.org	ie me, young-and-yummy doctor. </s><s> Can I	help	you? </s><s> I don't know. </s><s> Can you? <
30	foreverdreaming.org	know what? </s><s> Hey, you guys need some	help	? </s><s> Owen: where's webber? </s><s> You
31	foreverdreaming.org	er with her patients? </s><s> Now you need my	help	? </s><s> Don't push me, dude. </s><s> Okay, :
32	foreverdreaming.org	otta-you gotta get down here. </s><s> We need	help	. </s><s> Danny, Danny, I'm Dr. Hunt. </s><s> C
33	foreverdreaming.org	s><s> You've got one. </s><s> How can I be of	help	to you? </s><s> Sorry, I'm having a little bit of tr
34	foreverdreaming.org	stuff I got to get on top of. </s><s> So, you need	help	? </s><s> I need you to reschedule this meeting
35	foreverdreaming.org	><s> Help! </s><s> Please! </s><s> Someone	help	! </s><s> It's okay, buddy. </s><s> I got you. </s
36	foreverdreaming.org	> What are you doing? </s><s> We need some	help	in here. </s><s> We need to move the patient to
37	livejournal.com	n Jersey. </s><s> CHASE: [standing up] Can I	help	you? </s><s> This is a private room. </s><s> W
38	livejournal.com	ring] What happened? </s><s> Jessica: Please	help	me! </s><s> Jessica: Oh, my god! </s><s> [Car
39	livejournal.com	ther villager runs up.] </s><s> Villager #2: Help,	help	! </s><s> Dr. Sebastian, come quickly! </s><s>
40	livejournal.com	enters cheerily.] </s><s> CAMERON: How can I	help	you? </s><s> HOMELESS OLD GUY: The doct

Fig. V: Concordance lines for teaching set number 2 "Offering and asking for help" – Part II (SEMC)

For consolidation, the students can practice asking questions in a role-play. For this exercise, the learners can team up with their neighbor. The first step is to think of a typical situation they experienced on the ward which involved helping a patient, a family member or a colleague. Secondly, they are asked to write a short dialogue with the help of the concordance lines from the sheet. The next step is to assign the different parts and practise the role play. Once they feel confident enough and can speak freely, they act out their role play. This teaching method consolidates the competence of the learning group and constitutes a holistic approach by combining the introduction of a new term with practising and subsequently testing it (Legutke 2013: 92). The complete exercise sheet can downloaded by clicking on Fig. III-V.

3.3 Belly vs. stomach – appropriateness of word choice

This exercise is intended for the second year of schooling and targets the appropriateness of word choice. Cameron (1998: 209) places this particular language need in the field of 'Speech production accuracy'. With

regard to the guidelines of the Ministry of Education (NKM 2009a: n.p.) this worksheet can be used in class to address the competence areas 'Reception' and 'Production'.

As mentioned above, communication with patients is a core task of nurses. The communication of nursing professionals is characterized by a nursing care jargon and an interdisciplinary nature of the language in use (Marchwacka 2018: 173). Since the patient has little or no medical knowledge, the communication of nurses has to function as a bridge between the two opposites. Therefore, it contains a mixture of technical terminology, technical jargon, and everyday language (Marchwacka 2018: 179). Nonetheless, the nurses still have to maintain a professional distance to patients in their language. It is mandatory to differentiate between a factual and a relationship level, which makes the avoidance of colloquial expressions – especially with regard to excretions and particular body parts – absolutely essential (Menche 2007: 178). For this reason, the teaching material targets the appropriateness of word choice. A keyword analysis of the SEMC corpus was performed in order to identify frequently used colloquial expressions in conversations with patients and health professionals. This was followed by a keyword analysis in the MWC to determine the equivalent of the informal words.

A total of four colloquial expressions were chosen for the exercise, which are: *belly*, *poop* (v/n), *pee* (v/n), and *puke* (v/n). As a next step, a simple concordance analysis was performed to extract suitable sentences which display the relevant word in context. For each word, five sentences were chosen for the exercise sheet. With the exception of sentences containing the word *stomach*, all sentences were extracted from the SEMC corpus. The concordance search for the word *stomach* was not satisfactory due to decontextualization of the search results. Therefore, the sentences for *stomach* were selected from the MWC instead.

The assignments make the students use the corpus material and take a similar form to the exercises suggested by Gabrielatos (2005: n.p.). First, they have to find a German translation and decide whether the word is a noun or a verb. Then they have to determine whether the word is used formally or informally. Finally, they have to match the words, so that each informal expression has a matching formal equivalent. The exercise sheet can be viewed as Fig. VI-VII.

English / Communication	Accuracy of Word Choice	
Date:	Exercise Sheet	

Look carefully at the following sentences with examples of different nouns and verbs.

- What is the German translation of the word?
- Decide if the word is formal or informal.
- Is it used as a verb or a noun?
- Find the formal equivalent for each informal word.

1. Then we construct a pathway for a whole new **belly** button.
2. George (to Meredith): Look at her **belly**. </s><s> She's almost as wide as she is tall.
3. Unfortunately, she also needs her **belly** explored to evaluate the source of her internal bleeding.
4. Chase: Fever and **belly** pain. </s><s> Could be a gall bladder infection.
5. Raised her liver enzymes and caused the **bellyache**. </s><s> House: But not the fever.

1. Meredith: He'll be **urinating** out of a bag for a very, very long time.
2. Richard: Did you **urinate** in it? </s><s> Larry: No, I what?
3. It says here you haven't been able to **urinate** since yesterday. </s><s> Let's take a look.
4. Kayla: No. </s><s> Chase: Pain in **urinating**? </s><s> Kayla: No. 99's not that high, is it?
5. You have to get up ASAP after surgery. </s><s> Think you can **urinate**?

1. Oh and this morning I noticed my **poop** was a funny grayish color.
2. I fed him. </s><s> Oh god there's poop everywhere. </s><s> Wait. </s><s>
3. ... I won't go bald, get fat, and **poop** my pants.
4. And we're just gonna send her home farting and **pooping** uncontrollably for the rest of her life?
5. What if she was sloppy about washing her hands after **pooping**?

1. You don't believe me? </s><s> Punch me in the **stomach**, right now.
2. How long have you been having **stomach** pains?
3. Uh, he's really hot and his **stomach**'s hurting so much.
4. Arche Katz ... he came down with the **stomach** flu, and they asked me sub in.
5. Taub: How's your **stomach**? </s><s> Any diarrhea?

1. Mrs. Gibson: I have to **pee**. </s><s> Again. </s><s> (She sits up more and Mr. Gibson stands up to help her get out of bed)
2. Do you know he **peed** a little bit every time he coughed?
3. Like an incontinent old woman. </s><s> He smelled like **pee**. </s><s>
4. Paramedic: Ne, he **peed** all over himself! </s><s>
5. He looks down to see that the catheter has come out and he's **peed** his bed.

Fig. VI: Teaching set number 3 “Belly vs. stomach- appropriateness of word choice” – Part I (Vorbrink 2020: 72, concordance lines: SEMC; MWC)

1. Well, of course it is. </s><s> I mean, you don't **puke** blood if it's not serious.
2. House: Chase, I need you to head back to that "cesspool" and get me the kid's **puke**.
3. You didn't have strep or stomach bug the morning you came in, you **puke** because you were hung over.
4. I'm pretty sure Jack's **puke** is in a blue trash bag with the wrapped gifts, but if you find spaghetti, that's the wrong vomit.
5. Thank you and, as always, have a pleasant flight. </s><s> [House sits next to Cuddy.] </s><s> House: Still feel like **puke**ing?

1. If there are less than 5 loose **bowel movements** per day then dehydration or electrolyte imbalances are unlikely.
2. She had a **bowel movement** no more than once a week even though she consumed large amounts of laxatives.
3. Always clean up after a **bowel movement** with wet toilet paper.
4. Try to have three **bowel movements** a day.
5. He had a **bowel movement** while in the emergency room.

1. Do you think I'm pregnant? </s><s> Abby: No. </s><s> The **urine** test came back negative.
2. George: I noticed that **urine** output has increased slightly since yesterday.
3. I've been emptying **urine** bags all morning.
4. No, you'll pollute the sample! </s><s> Chase, get **urine** from the catheter.
5. He makes a face as the smell of the **urine** meets his nose.

1. Let's intubate! </s><s> Malucci: Not yet, I can bag her, all right? </s><s> Carol: She could **vomit** and aspirate.
2. [Lucy coughs again and blood spatters on page; Luke turns as she begins to **vomit** up lots of blood.]
3. Eric, lay down... his abdomen is rigid. </s><s> There's blood in his **vomit**
4. The amount of **vomiting** that kid did, would be enough toxins left in his system to show up in his blood.
5. Izzie: You're a liar. </s><s> And you also smell like **vomit**.

1. Diarrhea! </s><s> Blood in the **stool**! </s><s> These are routine questions--
2. [Foreman knocks and enters the room.] </s><s> Foreman: We need a **stool** sample.
3. Foreman: **Stool** samples were negative for parasites.
4. Split yourselves into 2 groups, test her blood, test her **stool**.
5. I asked what you did find in the **stool** sample. </s><s> Foreman: **Stool**. </s><s> And traces of iron, zinc, calcium carbonate, can we leave?

The teaching set can be further developed by watching short extracts of the episodes of the fictional medical TV series on the basis of which the SEMC corpus was compiled. As an observational task, the students could be asked to identify colloquial expressions in conversations they have watched and to rephrase the transcripts in appropriate medical language. This would ensure the repetition of the linguistic items, as required by Hutchinson/Waters (2010: 108). The teaching set would also improve the students' listening competence, as required by the Ministry of Education (NKM 2009a: n.p.).

3.4 Getting information – conducting a patient interview

According to Epp/Lewis (2009: 308), about 30 percent of the speech acts of nurses involve asking for information and giving instructions to patients or family members. In this teaching set, the focus is on the 'Clinical performance' of the nursing students. Cameron (1998: 211) distinguishes two different needs for the particular competence of 'Getting information', namely a medical interview (assessment) and a physical examination. The former is a task that is performed at the early stage of a patients' admission to the hospital on the ward. The assessment of patients' independence and their need of care are of special interest. In order to determine the individual needs, the nurses have to get the patients' medical history (Menche 2007: 302). The latter refers to the fact that an assessment usually involves a brief physical examination with nursing procedures such as measuring temperature, dressing a wound or giving medication (Oelke 2007: 639). This teaching set will exclusively focus on the patient interview and not on the physical examination. To my knowledge, there are no teaching materials that address the specific language task of getting information from a patient. In her textbook of basic vocabulary for nurses, Jibidar (2017: 151) presents a dialogue for a patient assessment but merely limits it to asking for a patient's phone number and address. Compared with the sample patient records which are used in Germany (cf. Oelke 2007: 638; Menche 2007: 308), it becomes clear that Jibidar's short extract is insufficient to meet the particular needs of German nursing students. The worksheet covers the competence areas of 'Interaction' and 'Production' (NKM 2009a: n.p.) and is intended for the third year of schooling (see table 2).

For the worksheet, keywords from a sample patient record based on Oelke (2007: 638) were extracted. The words were entered into Sketch Engine in order to perform a concordance search in the SEMC corpus. The extracted keywords are listed in the table below (table 3). With the help of the concordance lines, a language support sheet was developed for the nursing students. The support sheet is based on teaching materials designed by Grieser-Kindel/Henseler/Möller (2012: 106-107). It supports the students with a variety of ways to ask for a certain kind of information in the patient record. The language support sheet is designed to equip the students with enough necessary language knowledge, so that they can practise the act of conducting a patient interview in the subsequent class sessions (Hutchinson/Waters 2010: 109).

Table 3: Lemmas used for the concordance search in the SEMC corpus for teaching set number 4 (listed alphabetically and clustered by topic) (Vorbrink 2020: 44)

- Address
- Aids – Glasses – Hearing Aid – Cane – Walker
- Allergy
- Diet
- Documents
- Emergency Contact – Family
- Habit – Drinking – Smoking – Drugs
- Job
- Medication
- Mobility
- Sleep
- Symptoms – Problem – Illness

For Instruction / Information:

- Need to / Have to
- Hello

The phrases and sentences used in the language support sheet are exclusively derived from concordance lines from the SEMC corpus to ensure the use of authentic language. The support sheet is in Fig. VIII + IX. Additionally, an English patient history form was created (see Fig. XI) considering the listed information based on Oelke (2007: 638). The history sheet asks for basic patient information usually contained in every German assessment sheet. It is important to highlight that there is a clear separation between the medical assessment and the nursing assessment in Germany. The former is conducted by the medical staff and contains information regarding the clinical picture of the patient in conjunction with their medical history, such as their history of heart diseases (ibid.). Therefore, these types of questions are not included in the language support sheet for the worksheet, but are included in the medical history form for the sake of authenticity. It is important to note that the patient form is not given to the patients but instead exclusively filled out by the nurses (ibid.). Therefore, the students need to be trained how to ask the right questions in order to fill out the form correctly.

English / Communication	Getting Information	
Date:	Language Support Sheet	

Introduction	<p>-Good morning/ Hello, I'm</p> <p>-I need</p> <p style="padding-left: 40px;">-to ask you some questions.</p> <p style="padding-left: 40px;">-you to wait outside.</p> <p>-You must be very confused / in a lot of pain right now, but I need to ask you some questions, okay?</p>
Personal information	<p>-What's your permanent home address?</p> <p>-What is your job? What do you do for a living?</p> <p>-Who is your emergency contact? / Whom can I list as your emergency contact?</p> <p>-Do you have family nearby?</p> <p>-Do you have any paperwork, any transfer documents?</p>
Nursing-relevant information	<p>-Are you on a special diet?</p> <p>-How is your sleep pattern? / Do you have sleep problems?</p> <p>-Do you have unhealthy habits, such as drinking or smoking?</p> <p>-Do you have any allergies?</p>
Aids	<p>-Do you need / wear ...</p> <p style="padding-left: 40px;">-glasses / contact lenses?</p> <p style="padding-left: 40px;">-hearing aids?</p> <p>-Do you have mobility problems?</p> <p style="padding-left: 40px;">-Do you walk with the aid of a walker?</p> <p style="padding-left: 40px;">-Do you have / use a walking stick?</p>
Medication	<p>-I need to know if you are taking</p> <p style="padding-left: 40px;">-any medication.</p> <p style="padding-left: 40px;">-any prescription medication.</p>

Fig. VIII: Language support sheet as part of the teaching set number 4 – Part I (Vorbrink 2020: 74, based on Grieser-Kindel/Henseler/Möller 2012: 106-107; corpus: SEMC)

English / Communication	Getting Information	
Date:	Language Support Sheet	

	-sleeping pills. -antidepressants. -Are you allergic to any medication?
Current symptoms	-What are your symptoms? -What is your major symptom? -Do you have any other symptoms? / What are your other symptoms?
Saying goodbye	-Is there any family we should contact? / Is there anyone I can call for you? -Thank you! -That's all I need to know.

Reacting	-But we need to know the truth so we can help you. -You don't need to tell me. -We'll need to get your medical records.
-----------------	---

Fig. IX: Language support sheet as part of the teaching set number 4 – Part II (Vorbrink 2020: 75, based on Grieser-Kindel/Henseler/Möller 2012: 106-107; corpus: SEMC)

The assignments are marked by a three-way approach. Firstly, the students are asked to familiarize themselves with the patient history form and to look up unknown words in a dictionary. Since the main competence focus of the third year of schooling is on interaction and the communication skills of the learners (see table 1), the worksheet is designed to equip the students with sufficient language support so that they can fully engage in the communication task and benefit as much as possible from the time in class (Legutke 2013: 95).

English / Communication	Getting Information	
Date:	Exercise Sheet	

As a nurse, you are constantly asked to take a patient's medical history and get information from them. Today we will look at how you conduct a patient's admission in English.

1. Make yourself familiar with the language support sheet. Use a dictionary to look up words you don't know.
2. Now, turn to the patient history form. Which questions from the support sheet help you to get specific information for each section of the patients' medical history? Use colors to highlight the matching parts.
3. Pair up with your neighbor. Practise a fictional patient interview and take turns. Remember, you are practising the role of the nurse in this part.

Fig. X: Exercise sheet for teaching set number 4 (Vorbrink 2020: 76)

Next, the students are asked to highlight matching parts of the language support sheet with the patient history form to provide an orientation for the students. The last step is to team up and actually practise a fictional patient assessment. The exercise sheet can be downloaded by clicking on Fig. X. Legutke (2013: 97) points out the importance of students practising acquired competences. It enables the learners to prove their language skills and to experience the language in a real-case scenario (Legutke 2013: 98). Therefore, the worksheet, which is designed for a 90-minute classroom session, should also be used in the following lessons. According to the learners' proficiency levels, the following class sessions should offer additional practising time and eventually the actual presentation of the role play (Legutke 2013: 94). Getting the students to videotape their medical interview scenes could modify the act of presenting. Ideally, this would be done in the nursing school's skills lab. Rattani/Kaakour (2019: n.p.) encourage the use of self-produced screenplays in the classroom. Additionally, this innovative teaching method can be used to facilitate the students' production skills by focusing on the act of screenwriting instead of a more improvised acting approach, as evident in my worksheet. It also facilitates the students' media competence, as required by Trübswetter/Figueiredo (2019: 345).

PATIENT HISTORY FORM

PERSONAL INFORMATION: Patient's name: _____ Date of birth: _____ Address: _____ Phone number: _____	Date: _____ Name of nurse: _____ <input type="checkbox"/> Patient history compiled with a third party: _____ (name, surname)	PRIMARY DIAGNOSIS: _____ _____ _____ _____
Housing situation: <input type="checkbox"/> own household <input type="checkbox"/> in a household with: _____ Family members help in the household with: _____ _____	Size: _____ Weight: _____ BMI: _____ IMPORTANT INFORMATION: ALLERGIES: <input type="checkbox"/> yes <input type="checkbox"/> no _____ IMPORTANT MEDICATION: _____ _____	SECONDARY DIAGNOSIS / OPERATIONS: _____ _____ _____ _____
Employment (type): _____ Emergency contact: _____ (name, surname) _____ (phone number)	AIDS: <input type="checkbox"/> hearing aid left / right <input type="checkbox"/> glasses / lenses <input type="checkbox"/> walker / walking stick OTHERS: _____ CURRENT SYMPTOMS: _____ _____	FAMILY HISTORY: _____ _____ _____ DIET: <input type="checkbox"/> yes <input type="checkbox"/> no If yes, please specify: _____ HABITS: <input type="checkbox"/> non-smoker <input type="checkbox"/> smoker ADDICTION PROBLEM: <input type="checkbox"/> yes <input type="checkbox"/> no
Family doctor: _____ Transfer documents: <input type="checkbox"/> yes <input type="checkbox"/> no Native language: _____		

Fig. XI: Patient history form – Teaching set number 4 (Vorbrink 2020: 77, based on Oelke 2007: 638)

3.5 Additional contexts of *help* (v/n)

The last unit of the teaching sets developed in the course of the present chapter is designed for the third year of nursing schooling (see table 2). Unlike the other teaching materials, this exercise is characterized by the learners' direct use of the concordancing tool Sketch Engine and makes the students actively engage with the SEMC corpus. According to Cameron (1998: 214), the targeted language need is 'Speech production accuracy' in the context of 'Vocabulary learning' and the 'Clinical performance'. More precisely, it facilitates the ability of 'recognizing and deflecting cranky behavior and insults' in order to respond in a de-escalating way (ibid.). The relevant competence areas of the Ministry's guidelines are 'Interaction' and 'Production' (NKM 2009a: n.p.).

As mentioned above, Epp/Lewis (2009: 308) have pointed out that the speech acts of offering and asking for help are two of the main language skills of nurses. The teaching set presented in section 3.2 addresses this particular need. However, since the lemma *help* is highly frequent in the SEMC corpus (count 723), it is necessary to look at additional meanings of the word besides the ones identified by Epp/Lewis (ibid.). This should enable the students to accurately respond to statements by patients and family members involving the word *help* in a context other than asking for it.

Knowing a word involves various aspects, such as the written and spoken word form, knowledge of grammar, and knowledge of its meaning and associations (Nation 2013: 145). Nation (2013: 132) lists a range of activities that are suitable for vocabulary learning in order to address the meaning, concept, and reference of a word. My worksheet is based on Nation's suggestions to strengthen the concept and reference of a word. This particular aspect of vocabulary knowledge 'involves having a clear idea of the underlying meaning of a word that runs through its related uses, and also involves being aware of the range of

particular uses it has, that is, what it can refer to' (Nation 2013: 134-135). This activity enables the learner to understand the word *help* when it is used in a new context. Moreover, it helps the students to use the word in new ways (ibid.). Useful techniques to facilitate the named skills are 'finding common meanings' and 'playing at word detectives' (Nation 2013: 132). The former is a technique which helps the learners to understand the core meaning of a particular word (Nation 2013: 135). The latter refers to an activity where the learners look for information about a word they are already familiar with (Nation 2013: 136). The worksheet combines both activity types for the word *help* (v/n).

As mentioned above, in their previous year the students have been introduced to the word *help*. Naturally, the students are left with a positive connotation of the word, since they have intensely experienced it within that particular context. In this exercise, the students will be introduced to new uses of the word *help*.

- 1) Alex: Martin, we're trying to help.</s><s>Robert: I don't want your help.</s><s>(Alex
- 2) Derek: Look, I appreciate you're trying to help, but -</s><s> Other doctor in elevator:
- 3) This is not helping! </s><s>Meredith: She fell.</s><s>When she twisted her ankle, she
- 4) Alex: Martin, we're trying to help.</s><s>Robert: I don't want your help .</s><s>(Alex
- 5) that... that is not helping!</s><s>Why don't you tell me what is, then?</s><s>Talk to

Notice the differences of usage in example 3) and 5) in comparison to the usage in 4). Both have a negative connotation, but in 4), the person uttering the sentence rejects the help that is being offered, whereas in 3) and 5), the focus is on the fact that the help which has been offered and accepted is not helpful. Patients' statements like these require a different type of response from the nurses. This particular corpus-based teaching set is supposed to raise awareness of the versatile use of *help* in the medical work field in order to induce an appropriate reaction in response to the explicit meaning of the sentence.

It is important to note that in order for the students to work with a corpus compiled by their teacher, such as the SEMC, it is necessary for them to create an account beforehand. This way, the teacher can share the specialized corpus with the class. It is recommended to do this prior to the actual class session. Ideally, this can be combined with an introductory class session on Sketch Engine and corpus linguistics in general. An introductory session might contain basic information on the tool itself. Sketch Engine's website (Sketch Engine 2020: n.p.) provides a short YouTube video well-suited for this purpose and presents an additional listening comprehension for the students.

In this hard version of using corpora, the students have direct access to the teacher's SEMC corpus. Stevens (1991: 42) underlines the importance of the worksheet design for this type of teaching material. It has to walk the students through every single step of the process as closely as possible and gradually release them to their independent investigation of the concordance lines. The assignment sheet has been designed to meet that particular need by integrating screenshots from the homepage itself to underpin the directions (see Fig. XII-XIV). Both students and teacher will be provided with laptops from the school and a free version of Sketch Engine. Since each classroom is equipped with a projector, the teacher can demonstrate each step to ensure additional guidance.

English / Communication	Additional Contexts of <i>Help</i>	
Date:	Instructions	

During the past years of schooling, you have worked with different kinds of teaching materials derived from corpora. The presented concordances have been carefully selected for you by your teacher. Now it is time for *you* to explore a corpus and investigate patterns of language. Team up with your neighbor.

The following steps will guide you through the interface of Sketch Engine step by step. **Let's get started!**

Step 1) Start the laptop in front of you and open the browser.

Step 2) Type <https://www.sketchengine.eu> into the browser's address bar.

Step 3) Click on the **blue login button** on the Sketch Engine start page.



Step 4) Enter your **username** and **password**.

Step 5) Go to “**Select corpus**”.

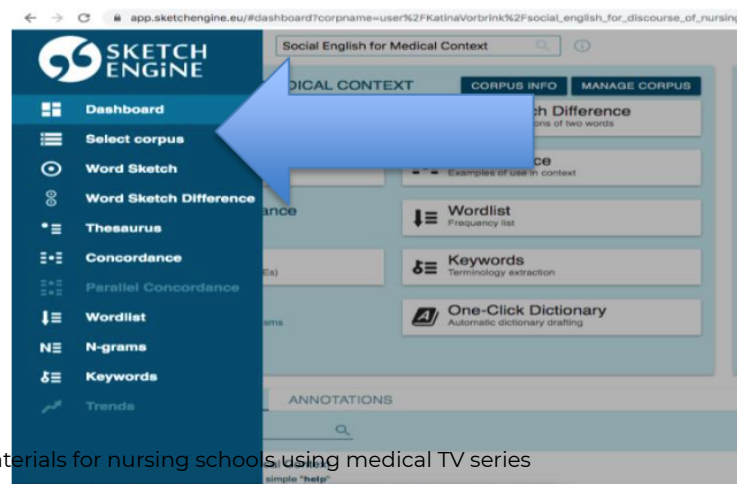
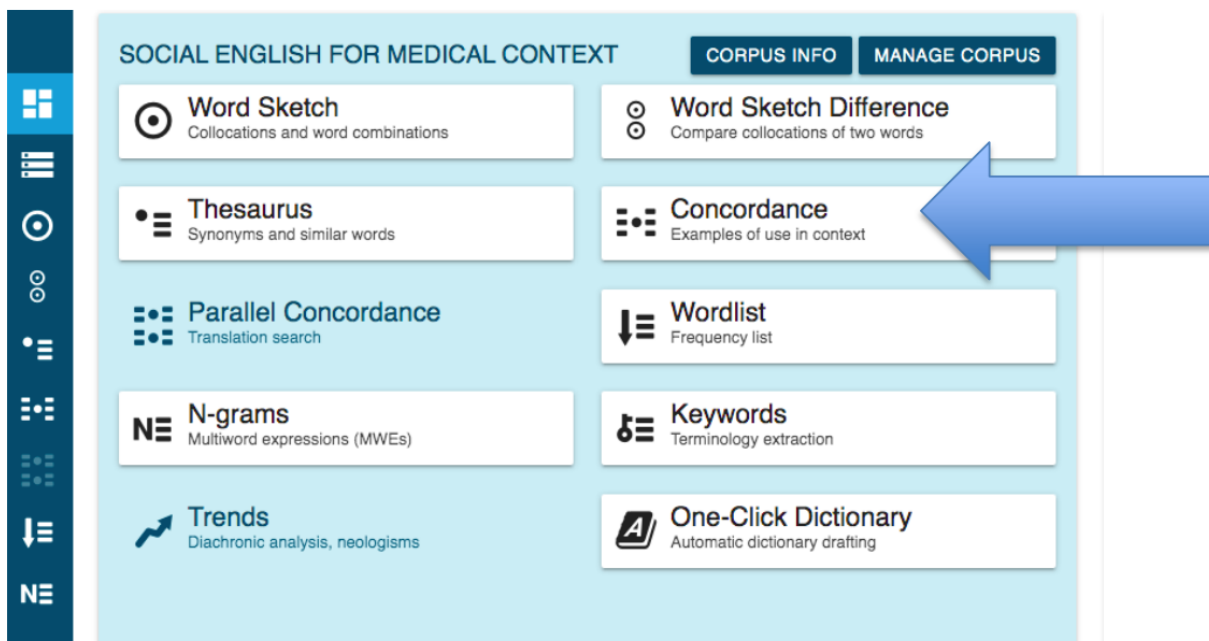


Fig. XII: Instruction sheet for teaching set number 5 – Part I (Vorbrink 2020: 78, screenshots: Sketch Engine 2020: n.p.)

Step 6) Now click on the shared corpus “**Social English for Medical Context**”.

Step 7) Click on “**Concordance**”.



Step 8) Click on the tab “**ADVANCED**” and type the word “**help**” into the search field.

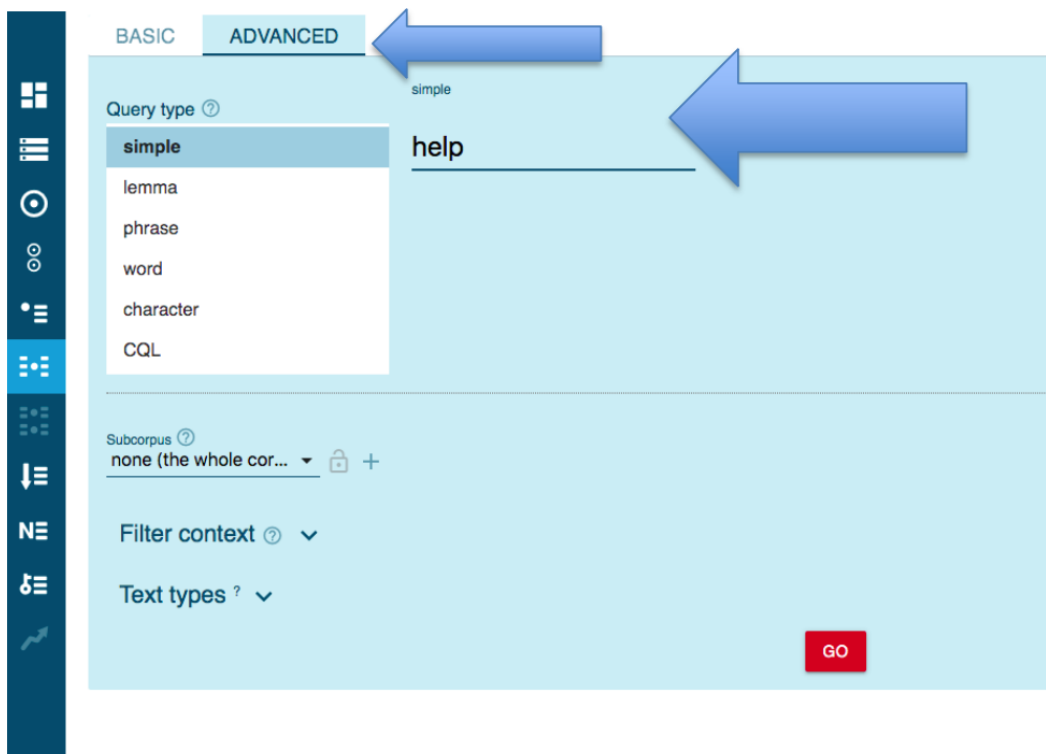
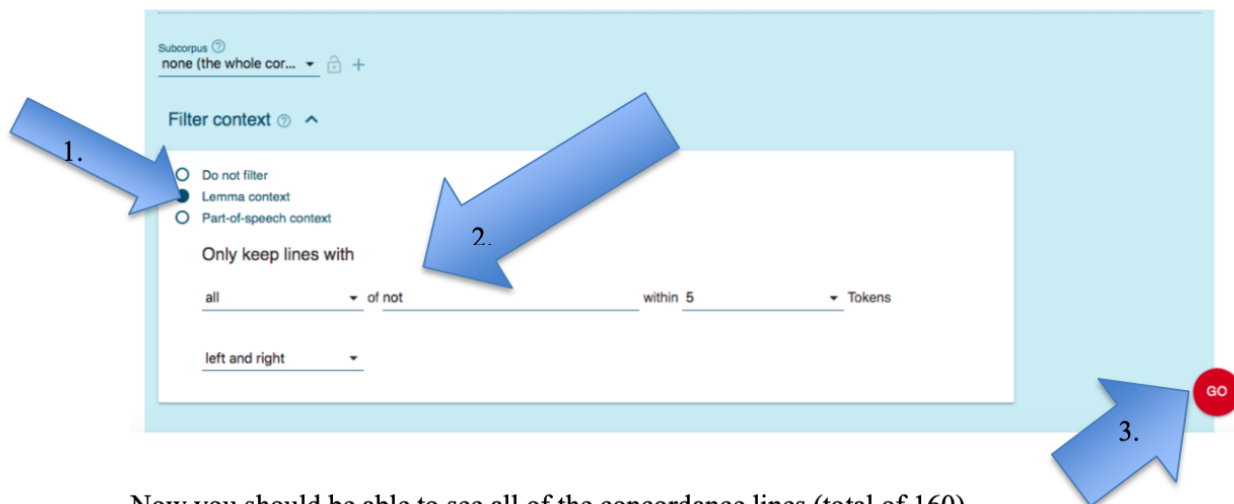


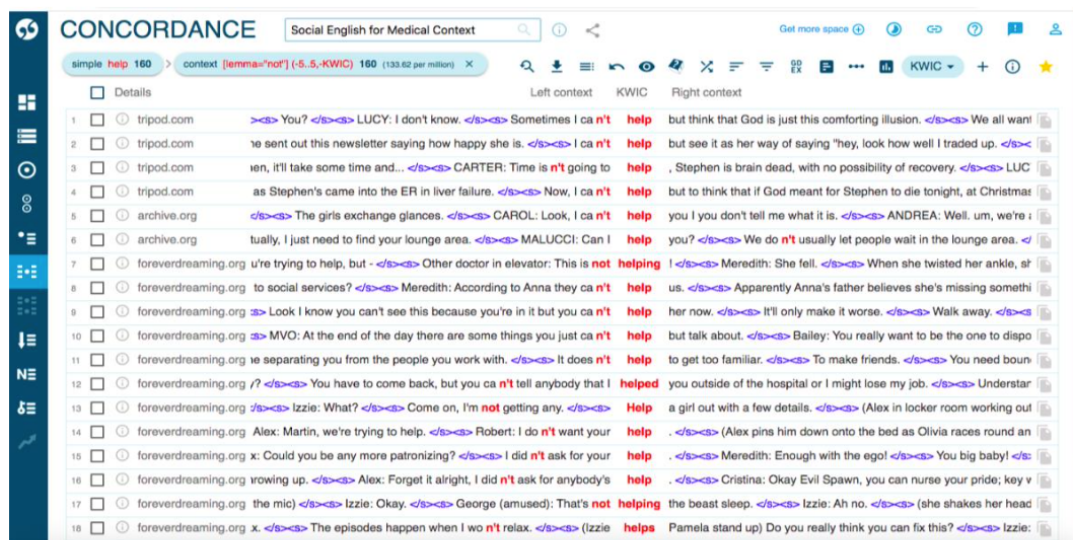
Fig. XIII: Instruction sheet for teaching set number 5 – Part II (Vorbrink 2020: 79, screenshots: Sketch Engine 2020: n.p.)

Step 9) Now select the filter “**Lemma context**” and type in the word “**not**”.

Press the red **go-button!**



Now you should be able to see all of the concordance lines (total of 160).



Assignments:

- Take a look at the lines and select the ones where *help* is used in a different way than offering help or asking for it. You can **tick the boxes** to mark the lines.
- Once you finished looking through all of the pages, you can look at your selection by clicking on the blue button “**SELECTED**” at the bottom of the page.
- Examine your selection closely. Think about how *help* in this context is different from the context of asking for help or offering it. Take notes.

Fig. XIV: Instruction sheet for teaching set number 5 – Part III (Vorbrink 2020: 80, screenshots: Sketch Engine 2020: n.p.)

The assignments are based on Gabrielatos’ (2005: n.p.) and Kennedy/Miceli’s (2001: 82) approaches to the hard version of corpus use in language teaching and learning. The first step is to introduce the students to the underlying question: ‘In what other ways can you use the word *help* other than to ask for it or offer it?’

This is followed by the outlined search strategy of asking the students to perform an advanced concordance search in the SEMC corpus with the lemma of interest. The added lemma context is *not* with five tokens for both sides left and right. The advanced concordance search of *help* in connection with *not* makes the overall number of concordance lines more manageable (count 160) for the students. In comparison, a simple concordance search gives an overall count of 887 lines. Asking the students to search all 887 instances might be overwhelming and it will be particularly difficult to identify additional patterns of the word of interest.

After the concordance search is completed, the students are asked to analyze the concordance lines for uses of *help* in connection with *not*. This marks the third step identified by Kennedy/Miceli (2001: 82) and is referred to as 'observing the examples found and selecting relevant ones'. During their search, students will come across familiar concordance lines used in the worksheet from the second year of schooling (see section 3.2). As mentioned above, it is very important that the teacher ensures close guidance of the students, for example by demonstrating how to see the word in fuller context by clicking on the key word (Gabrielatos 2005: n.p.). Also, Sketch Engine has a selection function, so that after the students have examined all concordance lines and ticked suitable ones for the particular context, they can look at their selected lines.

Once the students turn to their downsized choices, they are asked to examine the use of *help* in this new context. For further investigation, they can either print out their selection or work with a copy. They can also open a word document and paste the concordance lines into the document. After the students have classified the new forms, the next step, according to Thomas (2017: 20), is to form generalizations. According to Kennedy/Miceli (2001: 82), this marks the last step in the learners' investigation and is referred to as 'drawing conclusions'. One possibility to continue the teaching set in the following classroom sessions is to practise how to appropriately react to the diverse uses of *help* in the clinical setting.

4. Conclusion

The English language teaching materials developed in the course of the present study are a first step towards successfully implementing the subject 'English / Communication' in the new curriculum of nursing education in Germany. So far, the worksheets are merely theoretical and still need to be practically tested in the classroom setting.

Corpus-based teaching materials should be seen as an addition to existing teaching methods and materials (Gabrielatos 2005: n.p.). They are not at any time intended to replace conventional language-teaching methods, such as the work with textbooks.

Although a variety of aspects relevant for the design of corpus-based teaching materials was considered in the present study, a closer examination of other characteristics of nursing remains necessary to ensure the teaching materials' maximum usability. These include cultural differences and varying ways of understanding, communicating and coping with health-related issues (Bosher 2013: 265).

Another factor still in need of examination are the language needs of German nursing students. Since the care reform is fairly recent, we still lack a particular needs assessment in Germany with regard to English language skills in class and clinic. Therefore, the language needs for the teaching materials developed in the course of the present study were derived from Anglo-American and Canadian studies conducted by Cameron (1998) and Epp/Lewis (2009). Yet, these countries have different training systems. For this particular reason, it is quite likely that a German assessment will yield differing needs to varying degrees. Compared to other countries, the nursing profession in Germany is organized in two different ways: as an apprenticeship and as a course of studies. Regardless of the chosen course of education, the nurses perform similar tasks at work that do not differentiate with regard to the level of responsibility or level of payment

(Baumgarten/Ayerle 2016: 53). However, some fundamental differences could be identified. For example, unlike in Germany, academic skills such as writing texts and reading studies are part of the American nursing tasks (Cameron 1998). Furthermore, German nurses are predominantly responsible for therapeutic and nursing procedures, which do not include medical prescriptions or examinations, as is the case for their American colleagues. These differences can be attributed to the diverging education systems: in most countries, nursing training is almost exclusively located in the academic field and is not offered as dual training as it is in Germany (Menche 2007: 56-57). This stresses the necessity for a German needs assessment on the one hand, but also underlines restrictions that apply to transferring the foreign needs analysis to the development of teaching materials for German vocational schools on the other. Nevertheless, the foreign needs analysis offers a suitable starting point for determining the needs of German nurses with regard to English language learning in class and clinic.

Bosher (2013: 265) underlines the importance of cultural differences in nursing education because they can affect the communication with patients and their family members in various ways. It is essential to take these aspects into account when selecting specific topics in English lessons in order for nursing students to develop a competent and sensitive handling of it in class and clinic.

With regard to the corpus used for the development of teaching materials for the nursing school context, it may be worth considering compiling a spoken corpus based on authentic English interaction in the German clinical setting rather than on communication of fictional American medical television shows.

Römer (2009: 92) advises to create a database which contains useful corpus-based exercises for nursing education. Additionally, the database should list suitable corpora that teachers and students can turn to for this particular setting.

Generally, teaching education should also be targeted in the course of the reform. One option could be to refine university training programs by offering elective courses. These could train future teachers in various fields of vocational education and target relevant aspects, such as topic selection, task development, and the use of corpus linguistics in connection to vocational language teaching. Ideally, the elective courses could help to raise teachers' awareness of students' needs and train them to determine learners' needs in general. Corpus linguistics could be incorporated into the courses as a potential teaching method in vocational language teaching.

References

Akbari, Zahra (2016) "The study of EFL students' perceptions of their problems, needs and concerns over learning English: the case of MA paramedical students", In: *Procedia- Social and Behavioral Sciences* 232, 24-34.

Aston, Guy (ed.) (2001) *Learning with Corpora*. Bologna: CLUEB.

Baumgarten, Madlen/Ayerle, Gertrud M. (2016) "Wie begegnen Lehrkräfte an Pflegeschulen den heutigen Auszubildenden?", In: *PADUA* 11, 53-58.

Belcher, Diana D. (2006) "English for Specific Purposes: Teaching to Perceived Needs and Imagined Futures in Worlds of Work, Study, and Everyday Life", In: *TESOL QUARTERLY* 40 (1), 133-156.

Bosher, Susan (2013) "English for nursing", In: Paltridge, B./Starfield, S. (eds.) *The Handbook of English for Specific Purposes*. 1st Edition. Malden, Oxford: Wiley-Blackwell, 263-281.

Bundesministerium für Gesundheit (BMG) (2018) Pflegeberufegesetz. Web. 06.04.2020
<<https://www.bundesgesundheitsministerium.de/pflegeberufegesetz.html>>.

- Cameron, Richard (1998) "A language-focused needs analysis for ESL-speaking nursing students in class and clinic", In: *Foreign Language Annals* 31 (2), 203-218.
- Donath, Reinhard/Klemm, Uwe (2013) "Sprachhandlungskompetenzen entwickeln in multimodalen Lern- und Lebenswelten", In: Bach, G. (ed.). *Englischunterricht*. 5ed, Stuttgart: UTB, 121-147.
- Epp, Lucy/Lewis, Catherine (2009) "Innovation in language proficiency assessment: The Canadian English Language Benchmark Assessment for Nurses (ELBAN)", In: Bosher, S./Pharris, D. M. (eds.) *Transforming nursing education: The culturally inclusive environment*. New York: Springer Publishing Company, 285-310.
- Forever Dreaming Transcripts (2020) *Grey's Anatomy Transcripts*. Web. 29.04.2020
<<https://transcripts.foreverdreaming.org/viewforum.php?f=11&sid=6f903f4df216a3a0d50a8e01156fae02>>.
- Gabrielatos, Costas (2005) "Corpora and language teaching: Just a fling or wedding bells?", In: *The Electronic Journal for English as a Second Language* 8 (4). Web. 30.05.2020
<https://www.researchgate.net/publication/261708844_Corpora_and_language_teaching_Just_a_fling_or_wedding_bells>.
- Grieser-Kindel, Christin/Henseler, Roswitha/Möller, Stefan (2012) *Method Guide*. Methoden für einen kooperativen und individualisierenden Englischunterricht in den Klassen 5-12. Braunschweig: Westermann Gruppe.
- Grieser-Kindel, Christin/Henseler, Roswitha/Möller, Stefan (2017) *Method Guide*. Schüleraktivierende Methoden für den Englischunterricht in den Klassen 5-10. Braunschweig: Westermann Gruppe.
- Guhde, Jacqueline A. (2003) "English-as-a-second language (ESL) nursing students: Strategies for building verbal and written language skills", In: *Journal of Cultural Diversity* 10 (4), 113-117.
- Hamar, Christoph/Kuckeland, Heidi/Schneider, Kordula (2019) "Veränderungsprozesse an Pflegeschulen", In: *PADUA* 14 (3), 143-154.
- Hoffman, Beth L./Hoffman, Robert/Wessel, Charles B./Shensa, Ariel/Woods, Michelle S./Primack, Brian A. (2018) "Use of fictional medical television in health sciences education: A systematic review", In: *Health Sci Edu* 23, 201-216.
- Hutchinson, Tom/Waters, Alan (2010) *English for specific purposes. A learning-centered approach*. 25th Edition. Cambridge: University Press.
- Jibidar, Adèle (2017) *Grundwortschatz Englisch für Pflege- und Gesundheitsberufe*. 2. Edition. Vienna: Facultas Universitätsverlag.
- Jürgensen, Anke (2019) *Pflegehilfe und Pflegeassistenz- Ein Überblick über die landesrechtlichen Regelungen für die Ausbildung und den Beruf*. Leverkusen: BiBB. Web. 30.06.2020 <<https://www.bibb.de/dienst/veroeffentlichungen/de/publication/show/10155>>.
- Kennedy, Claire/Miceli, Tiziana (2001) "An evaluation of intermediate students' approaches to corpus investigation", In: *Language Learning and Technology* 5 (3), 77-90.
- Klitzing von, Waltraut/Stoll, Hansruedi/Trachsel, Edith/Aldorf, Kurt/Bernhard, Annelis/Eze, Germaine/Spirig, Rebecca (2007) "Fachenglisch für Pflegendende: Aus der Praxis für die Praxis", In: *Pflege* 20, 41-47.
- Krüger-Brand, Heike E. (2006) "Lernen von 'Emergency Room'- Medizinische, ethische und rechtliche Aspekte ausgewählter Fälle der Fernsehserie in der Diskussion", In: *Deutsches Ärzteblatt* 103 (37), 2350-2352.
- Legutke, Michael K. (2013) "Lernwelt Klassenzimmer: Szenarien für einen handlungsorientierten Fremdsprachenunterricht", In: Bach, G. (ed.). *Englischunterricht*. 5ed, Stuttgart: UTB, 91-120.

Leucht, Lisa (2013) *“Die Darstellung und Analyse von Erkrankungen in medizinischen Fernsehserien (House, Scrubs, ER) unter besonderer Betrachtung von Nierenerkrankungen“*, Dissertation, Medizinische Fakultät Jena. Web. 25.04.2020 <<https://d-nb.info/1065233833/34>>.

Levshina, Natalia (2015) *“Online film subtitles as a corpus: An n-gram approach.“* Web. 12.04.2020 <https://www.researchgate.net/publication/286878971_Online_film_subtitles_as_a_corpus_An_n-gram_approach>.

LIVEJOURNAL Clinic-Duty (2020) *Transcripts Dr. House*. Web. 29.04.2020 <<https://clinic-duty.livejournal.com/12225.html>>.

Marchwacka, Maria A. (2018) *“Zur Kommunikation im Gesundheitswesen am Beispiel der pflegerischen Fachsprache – eine Bestandsaufnahme“*, In: Tinnefeld, T. (ed). *Fremdsprachenunterricht im 21. Jahrhundert: Lerner-Methoden-Herausforderungen*. Band 8. Saarbrücken: Saarbrücker Schriften zu Linguistik und Fremdsprachendidaktik (SSLF).

Menche, Nicole (2007) *Pflege Heute- Kommunikation in der Pflege*. 4. Edition. München: Urban & Fischer.

Minister Konferenz (MK) (2020) *“Informationen zur Neuordnung der Pflegeausbildung (Umsetzung in Niedersachsen)“* Web. 14.06.2020 <https://www.mk.niedersachsen.de/download/151031/Fragen_und_Antworten_zur_Generalistik_01.12.2020.pdf>.

Miyake, Misuzu/Tremarco, John (2005) *“Needs analysis for nursing students utilizing questionnaires and interviews“*, In: *Journal of Medical Welfare* 11 (1), 23-34.

Nation, I. S. P. (2013) *Learning Vocabulary in Another Language*. 2nd Edition. New York: Cambridge University Press.

Niedersächsischer Bildungsserver (NiBiS) (2020) *“Englisch im Sekundarbereich II“* Web. 02.07.2020 <https://www.nibis.de/englisch-im-sekundarbereich-ii_4312>.

Niedersächsisches Kultusministerium (NKM) (2009a) *“Rahmenrichtlinien für das Unterrichtsfach Englisch/ Kommunikation im berufsübergreifenden Lernbereich der Berufsfachschule, die auf dem Sekundarabschluss I – Realschulabschluss- aufbaut.“* Web. 31.01.2020 <https://www.nibis.de/nli1/bbs/archiv/rahmenrichtlinien/eng_bfs_II_dr.pdf>.

Niedersächsisches Kultusministerium (NKM) (2009b) *“Ergänzende Bestimmungen für das berufsbildende Schulwesen (EB-BbS).“* Web. 06.04.2020 <<http://www.nds-voris.de/jportal/portal/t/7mt/page/bsvorisprod.psml;jsessionid=A64B5D24D3FE4B4338375DBF84033DF5.jp10;jsessionid=26A36845463C289AC52B94DC15CDC273.jp10?doc.hl=1&doc.id=VVND-VVND000041157&documentnumber=1&numberofresults=2&doctype=vvnd&showdoccase=1&doc.part=F¶mfromHL=true#focuspoint>>.

Oelke, Uta (2007) *In guten Händen. Gesundheits- und Krankenpflege*. Berlin: Cornelsen Verlag.

O’Keefe, Anne/McCarthy, Michael/Carter, Ronald (2007) *From corpus to classroom. Language use and language teaching*. Cambridge: University Press.

Oswald, Birgit (2009) *“Expansion im Halbgötterolymp. Entwicklungen, Erfolgsfaktoren und Nutzungsmotive von Arztserien“*, Diplomarbeit. Universität Wien. Web. 13.04.2020 <<https://core.ac.uk/download/pdf/11588141.pdf>>.

Pflegeberufereformgesetz (PflBRefG) (2017) *“Bundesgesetzblatt Teil 1, Nr. 49.“* Bonn. Web. 30.06.2020 <https://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl117s2581.pdf#_bgbl_%2F%2F%5B%40attr_id%3D%27bgbl117s2581.pdf%27%5D__1593532298419>.

Rattani, Abbas/Kaakour, Abdul-Hadi (2019) "Screenplays and screenwriting as an innovative teaching tool in medical ethics education", In: *Journal of Medical Humanities*. Web. 07.04.2020. <<https://link.springer.com/article/10.1007/s10912-019-09584-8>>.

Römer, Ute (2009) "*Corpus research and practice. What help do teachers need and what can we offer?*" Web. 12.04.2020 <https://www.academia.edu/2831291/Corpus_research_and_practice>.

Savova, Lilia (2018) "Conversation analysis criteria for evaluating the authenticity of ESL textbook conversations", In: Azarnoosh, M./Zeraatpische, M./Faravani, A./Kargozari, H. R. (eds.) *Issues in coursebook evaluation series: critical new literacies: the praxis of English language teaching and learning* 8. Brill NV: The Netherlands.

Sketch Engine (2020). Web. 15.06.2020 <<https://www.sketchengine.eu/>>.

Simply Scripts (2006) *Online Scripts E.R.* Web. 29.05.2020 <https://www.simplyscripts.com/tv_ef.html#E>.

Stevens, Vance (1991) "Classroom concordancing: Vocabulary materials derived from relevant, authentic text", In: *English for Specific Purposes* 10, 35-46.

Thomas, James (2017) *Discovering English with Sketch Engine – A Corpus-Based Approach to Language Exploration*. 2nd Edition. Versatile.

Trübswetter, Angelika/Figueiredo, Lina (2019) "Digitalisierung in der deutschen Pflegeausbildung: Potenziale und Herausforderungen des AKOLEP-Projekts", In: *Pflege-Hogrefe* 32 (6), 343-352.

Vorbrink, Katina (2020). *Developing English for Specific Purposes Teaching Materials for the Curriculum of a Nursing School in Lower Saxony – A Corpus-Based Approach* (Unpublished Master's thesis, Osnabrück University, Osnabrück, Germany).

Wong, Roger Y./Saber, Sadra S./Ma, Irene/Roberts, Mark J. (2009) "Using television shows to teach communication skills in internal medicine residency", In: *BMC Medical Education* 9 (9), 1-8.

1 The German term is 'Pflegeberufereformgesetz' (PflBRefG). The new job title is 'Pflegefachfrau / Pflegefachmann'.

2 This subdivision of vocational schooling is typical of Germany and not common in other countries (Jürgensen 2019: 7).

3 The practical training takes place in different clinical settings, e.g., different specialized medical department as well as outpatient treatment and home care (PflBRefG 2017: 2584).

4 The specialized lessons are now thematically divided into 11 curricular units and deliberately distance themselves from the previous subject canon, i.e., teaching topics isolated from one another and not with underlying occupational situations (MK 2020: 4).

5 The first author has been working at the nursing school since 2017 and has been the head of curriculum development since December 2019.

6 The German term is 'Binnendifferenzierung'.

7 The German term is 'Pflegehelfer'.

8 The context in the concordance lines was not sufficient to grasp the meaning of the word.

9 This meets the general requirements defined in the guidelines from the Ministry of Education (NKM 2009a: n.p.), which focus on the ability of the students to use a dictionary correctly. However, this aspect of language learning is not addressed in the present chapter.

10 Source: SEMC corpus – concordances from an advanced concordance search with *help* and *not*.

11 While watching, the students can be asked to note down important information, so that they can give a short summary of the video.

12 Each school laptop has a direct connection to the general printer.

13 Further fields of vocational training in Germany include engineering, health sciences, and cosmetology.

Appendix: Additional resources

ELEN LE FOLL

This section is simply a list of additional, useful resources for you to continue your corpus linguistics journey.

This list aims to be dynamic: should you find a dead link, or that a resource is missing, do please send me an e-mail. In the spirit of this Open Educational Resource, freely available resources will be prioritised.

1 Online corpus tools and corpora

- English-corpora.org
 - Free basic access to all users and corpora (with a limited, but usually sufficient, number of queries).
 - Registration required.
 - Option to upgrade to premium individual or academic licenses.
 - Gives access to a range of English corpora covering general British and American English, web English registers, online news, Wikipedia articles, World Englishes, etc.
- Sketch Engine
 - Limited free trial account (30 days).
 - Registration required.
 - Many EU higher education institutions have access via the EU-funded ELEXIS project.
 - Gives access to a range of English corpora including the Spoken BNC2014, the EnTenTen15, the Cambridge Open Learner Corpus, many non-English corpora, and the possibility to build DIY corpora.
 - A number of open corpora are available without an account.
- SKELL (Sketch Engine for language learning)
 - Free and no registration required.
- CQPweb
 - Free
 - Registration required.
 - Gives access to several useful English corpora including the Spoken BNC2014, the American and British English 2006 corpora, and many corpora in languages other than English (e.g. Arabic, Norwegian, Punjabi, Thai).
- BNClab
 - Free and no registration required.
 - A learner-friendly interface to explore the spoken sections of the British National Corpora (BNC) 1994 and 2014.
 - The Corpus for School project website features many worksheets and lessons ideas to work with the BNClab.
- TCSE Ted Corpus SearchEngine
 - Free and no registration required.
 - A user-friendly search engine designed to explore TED talks with the transcripts and original videos.
 - Includes useful concordancing, n-grams and construction functions.
- Compleat Lexical Tutor
 - Free and no registration required.

- Includes a range of useful tools for data-driven learning.
- Michigan Corpus of Academic Spoken English (MICASE)
 - Free and no registration required.
- Michigan Corpus of Upper-Level Student Papers (MICUSP)
 - Free and no registration required.
- VersaText
 - Free and no registration required.
 - Allows teachers or learners to analyse individual texts.
- Leeds collection of Internet corpora
 - Free and no registration required.
 - Includes a range of Internet language corpora, including in languages other than English.

2 Free downloadable corpus tools for offline corpus work

- AntConc
 - Free to download and use for non-commercial purposes.
 - Works cross-platform (Windows, Mac OS X, Linux).
 - Can be combined with a whole suite of more specialised free software also developed by Laurence Anthony (e.g. AntCorGen to create discipline-specific corpora on the basis of open-access academic articles).
- LancsBox
 - Free to download and use for non-commercial purposes.
 - Works cross-platform (Windows, Mac OS X, Linux).

A far more comprehensive list of corpus tools can be found [here](#).

3 Additional free online resources for teachers and learners

- The Corpus for School website features many worksheets and lessons ideas based on the BNClab (see above) for A-level English and English as a Foreign Language (EFL) (project leader: **Dana Gablasova**).
- Improving writing through corpora: Data-driven learning? A free online course in five modules from the **University of Queensland** (also available in Mandarin Chinese)
- Datengesteuertes Lernen: Lexiko-grammatische Kenntnisse erweitern und sprachliche Zweifelsfälle lösen. An accessible article on data-driven learning with examples and worksheets for the EFL classroom by **Dirk Siepmann** (in German).
- Recordings of **Peter Crosthwaite**'s online workshops on corpora and data-driven learning for English Language Teaching (ELT) lesson planning.
- Video recordings of talks and/or slides on the use of corpora to inform the teaching of English for Academic Purposes from a one-day event organised by the **BAAL Corpus Linguistics SIG** and held online on 14 Jan 2021.
- A series of short video tutorials on using english-corpora.org and Sketch Engine by **Elen Le Foll** (recorded in 2020 and 2021).

4 Recommended books

- Bennet, Gena R. (2010). *Using Corpora in the Language Learning Classroom. Corpus Linguistics for Teachers*. University of Michigan Press: Michigan ELT.
- Crawford, William J., & Csomay, Eniko (2016). *Doing corpus linguistics*. New York: Routledge.
- Crosthwaite, Peter (Ed.). (2020). *Data-driven learning for the next generation: Corpora and DDL for pre-tertiary learners*. London; New York: Routledge.
- Friginal, Eric (2018). *Corpus linguistics for English teachers: New tools, online resources, and classroom activities*. New York: Routledge.
- Karpenko-Seccombe, Tatyana (2020). *Academic writing with corpora: A resource book for data-driven learning*. London: Routledge.
- O’Keeffe, Anne, McCarthy, Mike, & Carter, Ronald (2007). *From corpus to classroom: Language use and language teaching*. Cambridge: Cambridge University Press.
- Pérez-Paredes, Pascual (2020). *Corpus Linguistics for Education: A Guide for Research* (1st ed.). Abingdon, Oxon; New York: Routledge.
- Poole, Robert (2018). *A guide to using corpora for English language learners*. Edinburgh: Edinburgh University Press.
- Timmis, Immis (2015). *Corpus linguistics for ELT: Research and practice*. London; New York: Routledge.