



Universität des Saarlandes

Philosophische Fakultät (P)

Fachrichtung Anglistik & Amerikanistik

Dissertation

zur Erlangung des akademischen Grades
eines Doktors der Philosophie (Dr. phil.)

Corpora and data-driven learning in foreign language teaching

**Using a corpus of international video-mediated communication to develop
communicative and intercultural competence in the English as a foreign
language classroom**

Vorgelegt von:

Caroline Elisabeth Collet

Saarbrücken, 8. Oktober 2024

Dissertation zur Erlangung des Grades eines Doktors der Philosophie der Philosophischen Fakultät der Universität des Saarlandes

Dekanin der Philosophischen Fakultät: Prof. Dr. Nine Miedema

Erstgutachter: Prof. Dr. Stefan Diemer

Zweitgutachterin: Prof. Dr. Astrid M. Fellner

Tag der letzten Promotionsleistung: 26.06.2025

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	VI
ABSTRACT.....	VIII
LIST OF FIGURES.....	IX
LIST OF TABLES	XII
LIST OF ABBREVIATIONS.....	XIV
1. INTRODUCTION	1
2. RESEARCH BACKGROUND	10
2.1. ASPECTS OF TEACHING	11
2.1.1. <i>Didactics and dimensions</i>	12
2.1.2. <i>Foreign language teaching</i>	21
2.1.3. <i>English as a foreign language teaching</i>	23
2.1.3.1. English native and (non-native) speakers	23
2.1.3.2. Reference disciplines of EFL teaching	28
2.1.3.3. Development of EFL teaching in Germany.....	30
2.1.3.4. Levels of teaching.....	31
2.1.4. <i>Frameworks and standards</i>	32
2.1.4.1. Common European Framework of Reference for Languages (CEFR)	33
2.1.4.2. Education standards in Germany	41
2.1.5. <i>Competences</i>	44
2.1.5.1. Communicative competence	47
2.1.5.1.1. Communicative competence in the CEFR	49
2.1.5.1.2. Functional communicative competence in the German education standards	54
2.1.5.2. Intercultural competence	57
2.1.5.2.1. What is culture?	57
2.1.5.2.2. Aspects of intercultural competence	66
2.1.5.2.3. Byram's model of intercultural communicative competence	69
2.1.5.2.4. Intercultural competence in the CEFR.....	73
2.1.5.2.5. Intercultural competence in the German education standards	83
2.1.5.2.6. Looking at German federal state curricula	86

2.1.5.2.7. The intercultural speaker as the (new) model in EFL teaching?	98
2.1.6. EFL teaching and learning in the digital era	106
2.1.6.1. From digitalization to the culture of digitality	106
2.1.6.2. 21 st century skills and digital competences	112
2.1.6.3. Education in the digital era in Germany	114
2.2. CORPUS LINGUISTICS AND THE ROLE OF CORPORA IN EFL TEACHING	121
2.2.1. Foundations of corpus linguistics	121
2.2.1.1. What is a corpus?	121
2.2.1.2. What is corpus linguistics?	128
2.2.1.3. How to use corpora	132
2.2.2. Using corpora in EFL teaching	136
2.2.2.1. The authenticity gap	137
2.2.2.2. Influence of corpus linguistics on language teaching	139
2.2.2.3. Data-driven learning	141
2.2.2.4. How corpora can be used in the classroom	144
3. DATA AND METHODOLOGY	153
3.1. DATA	153
3.1.1. TaCoCASE, ViMELF and the CASE project	153
3.1.2. TaCoCASE	156
3.1.2.1. Corpus design	159
3.1.2.2. Corpus information	163
3.1.2.2.1. General numbers	164
3.1.2.2.2. Conversations and topics	166
3.1.2.2.3. Participants	169
3.1.2.2.4. Setting	173
3.1.2.3. Compilation process	179
3.1.2.3.1. Recording phase	180
3.1.2.3.2. Transcription phase	184
3.1.2.3.3. Anonymization and correction phase	190
3.1.2.3.4. Issues and challenges	191

3.1.3. <i>Benefits of TaCoCASE</i>	194
3.2. METHODOLOGY	196
4. ANALYSIS: LINGUISTIC AND (INTER)CULTURAL ASPECTS IN TACOCASE.....	213
4.1. LINGUISTIC ANALYSIS	213
4.1.1. <i>Frequency analysis</i>	219
4.1.2. <i>Lexical analysis</i>	226
4.1.3. <i>Collocations and n-grams</i>	238
4.1.4. <i>Discourse markers</i>	251
4.2. (INTER)CULTURAL ANALYSIS	278
4.2.1. <i>Cultural aspects in TaCoCASE</i>	283
4.2.1.1. Food	291
4.2.1.2. Customs and traditions	295
4.2.1.3. Popular culture – TV series and movies	300
4.2.2. <i>Talking about cultural aspects</i>	308
5. APPLICATION: TACOCASE AND EFL TEACHING	334
5.1. TACOCASE AND COMMUNICATIVE COMPETENCE	336
5.1.1. <i>Improving linguistic competence</i>	338
5.1.2. <i>Improving pragmatic competence</i>	351
5.2. TACOCASE AND COMMUNICATIVE LANGUAGE STRATEGIES	355
5.2.1. <i>Interaction strategies</i>	356
5.2.2. <i>Mediation strategies</i>	364
5.3. TACOCASE AND INTERCULTURAL COMPETENCE	374
5.4. MULTIMODAL FEATURES OF TACOCASE	386
5.4.1. <i>Online interaction</i>	388
5.4.2. <i>Non-verbal communication</i>	392
5.5. CASE STUDIES: HOW TACOCASE CAN BE USED IN THE EFL CLASSROOM	399
5.5.1. <i>How to teach vocabulary with TaCoCASE</i>	400
5.5.2. <i>How to explain (new) cultural concepts with TaCoCASE</i>	407

5.6. BENEFITS OF TACoCASE IN A TEACHING CONTEXT	416
6. CONCLUSION	421
REFERENCES	430
APPENDIX	451

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, **Prof. Dr. Stefan Diemer**, for the unwavering support, invaluable advice, and steadfast mentorship. Thank you for tackling this dissertation project together with me, granting me freedom to pursue tasks outside academia, while always encouraging me to continue my research and this project with determination. Your guidance throughout my entire journey, your invitation to be part of the CASE project, and the opportunities to attend conferences from the early stages of my career have provided me with invaluable insights into the worlds of both academia and applied linguistics.

I also wish to extend my heartfelt thanks to **Prof. Dr. Astrid Fellner**, who graciously agreed to co-supervise this dissertation and who supported me on so many levels. We share the love and passion for American culture and have worked together very trustworthily not only in academia, but also at the Deutsch-Amerikanisches Institut Saarland e.V., where we literally climbed mountains together. Thank you for being an inspiring mentor and a true role model.

I am immensely grateful to the **Konrad-Adenauer foundation** for providing me with a doctoral grant and for both the financial and idealistic support. A special thanks goes to **Dr. Susanna Schmidt, Dr. Michael Schmitz, Dr. Daniela Tandecki, Dr. Helena Köhler** for believing in me and offering me amazing opportunities: From selecting me to be a member of the Promovierendeninitiative e.V., offering me a spot in the Cusanus Karriereförderprogramm für Frauen and even sending me to meet with Bundespräsident Frank-Walter Steinmeier as a representative of the KAS scholarship holders. I would like to thank all fellow scholarship recipients for the enriching discussions during seminars and for becoming not only colleagues but true friends. I would like to thank my alma mater, **Saarland University**, for their financial support through the **GradUS program**, which allowed me to attend conferences and present my research to distinguished scholars as well as going on a research visit to Birmingham City University to meet up with our CASE project partners and to collect the data for the corpus.

This dissertation would not have been possible without the help and input from many individuals. My heartfelt thanks go to my CASE project partners in crime: **Dr. Marie-Louise Brunner** and **Dr. Selina Schmidt**. Thank you for sharing this academic journey, for many conference visits, the shared laughter, your valuable feedback and unwavering support. You have been wonderful colleagues and above all amazing friends. A huge thanks is due to all **participants and transcribers of TaCoCASE**, who made a significant contribution to the existence of the corpus and to the data that is the basis for this dissertation. A special thank you to **Prof. Dr. Sebastian Gerling, LL.M.** for reigniting my academic passion and ambition and for the final push that led me to complete this dissertation.

Thank you to **Prof. Dr. Andrew Kehoe and Matt Gee**, our CASE project partners from Birmingham City University, for your continuous support during the compilation and publication of TaCoCASE. We had a great time when I was on my research visit at BCU and at several conferences, where we shared thoughts, laughter and (traditional) food – fun and enlightening memories that I will cherish. And thank you to **Dr. Casey Iezzi** (Keck), our CASE project partner from Boise State University, who not only agreed to collaborate on this project and collected data from the Boise participants but who above all sparked my passion for corpus linguistics during her class I attended at BSU during my semester

abroad. Since then, corpus linguistics and its application in language teaching has captivated me and has led me to conduct my own research in this field. Thank you for being a true inspiration.

My friends, **Verena B., Monique, Verena M. and Sophie**, and my godchild **Emilia**, have been a constant source of support throughout this journey. Thank you for believing in me, pushing me beyond my limits, cheering me on and lending me an ear whenever I needed it. I regret the times I had to postpone or cancel meetings, but knowing all of you by my side and knowing that I can count on you gave me strength to push through this.

Last but not least I would like to thank my family. **My grandparents, Agnes and Erwin Collet**, who had a great impact on me when I was a child, who ignited my ambitiousness and encouraged me throughout my life to pursue my dreams. Though they did not live to see the completion of this dissertation, I know they would be immensely proud of me today. To **my parents, Oranne and Eike Collet, and my brother Philipp Collet**: No words can express how grateful I am for your unwavering support – not only during this dissertation but throughout my entire life. I owe everything I am to you. You never lost faith in me and my skills even when I doubted myself. Your unconditional love, encouragement and understanding gave me strong roots so that I can grow and stand against every storm.

Thank you!

ABSTRACT

The education system and English as a Foreign Language teaching (EFL) in particular should prepare learners to live and work in a globalized, intercultural and digital environment. Future skills like communication, collaboration, creativity and critical thinking are needed, and establishing communicative and intercultural competence, as well as digital competence in the EFL classroom is key.

This dissertation explores the potential of using corpora and data-driven learning as powerful tools to establish these competences and at the same time to modernize and upgrade language teaching. However, the use of corpora in EFL teaching is currently very limited and it seems that corpora do not find their way into EFL teaching and into the EFL classroom. To address this gap, this dissertation introduces new material and new methods to promote a new mindset towards integrating corpora and data-driven learning (DDL) in EFL teaching.

A key contribution of this dissertation is the creation of a new resource, TaCoCASE, a computer-mediated communication (CMC) corpus that consists of intercultural encounters between students from Germany, Great Britain and the United States. It provides authentic material and insights into actual language in use and the negotiation process of cultural concepts in intercultural encounters. This dissertation investigates to what extent corpora and DDL can contribute to improving the main competences in EFL teaching – communicative competence and intercultural competence, as well as digital competence. It encourages the use of corpora and data-driven learning in language teaching and has several significant implications for EFL teaching: (1) It demonstrates that there is a clear syllabus fit of TaCoCASE with European standards – the Common European Framework of Reference for Languages (CEFR) in terms of communicative and intercultural competence, (2) it shows how authentic examples and DDL activities can be used to develop communicative and intercultural competence.

Thus, this dissertation argues for a paradigm shift in EFL teaching through using corpora and DDL activities in the EFL classroom. This ranges from suggesting new material, new methods and a new mindset to practical applications with hands-on material and examples ready for classroom use.

LIST OF FIGURES

Figure 1: Didactic triangle	13
Figure 2: Educational objectives (Bloom 1956).....	15
Figure 3: Didactic square.....	20
Figure 4: EFL teaching, FLT and didactics	22
Figure 5: Kachru's circles of English.....	24
Figure 6: English didactics and its influences (Haß 2006: 10).....	29
Figure 7: Factors of language proficiency.....	35
Figure 8: CEFR reference levels	36
Figure 9: Updates of descriptors (Screenshot from CEFR 2020: 23)	39
Figure 10: From CEFR to classroom.....	42
Figure 11: Competences according to the German education standards (KMK 2012: 12)	43
Figure 12: Aspects of communicative language competence in the CEFR.....	49
Figure 13: Communicative language competences in the CEFR (CEFR 2001: 108f).....	50
Figure 14: Linguistic resources	56
Figure 15: Factors in intercultural communication (Byram 1997: 34)	70
Figure 16: Plurilingual and pluricultural competence	75
Figure 17: Example from FREPA database (screenshot FREPA)	82
Figure 18: Intercultural communicative competence (KMK 2012: 19)	84
Figure 19: SAMR-model (adapted from PuenteDura 2006)	109
Figure 20: TPACK-model (adapted from Koehler et al. 2014: 103)	113
Figure 21: Screenshot from AntConc (version 4.1.0.)	133
Figure 22: Applications of corpora in the classroom (adapted from Mukherjee 2002: 160)	144
Figure 23: Sub-corpora of the CASE project: ViMELF and TaCoCASE	154
Figure 24: Screenshot from WebCorp (I)	157
Figure 25: Screenshot from WebCorp (II)	158
Figure 26: Conversations – NS and NNS in TaCoCASE.....	165
Figure 27: Number of words per speaker combination	168
Figure 28: Conversation length per speaker combination	168

Figure 29: Gender distribution in TaCoCASE	169
Figure 30 Age distribution in TaCoCASE	170
Figure 31: Compilation phases	179
Figure 32: Screenshot of transcription software.....	185
Figure 33: Screenshot from conversation 06SB157BO11	190
Figure 34: Screenshot from conversation 07SB149BI31	190
Figure 35: Research questions for linguistic and (inter)cultural analysis	202
Figure 36: TaCoCASE application process	207
Figure 37: Overview of selected competences and objectives	208
Figure 38: Lexical density (mean).....	234
Figure 39: Type-token ratio (TTR) for different speaker groups	236
Figure 40: Frequent n-grams in spoken discourse (adapted from Biber et al. 2002: 444).....	240
Figure 41: Distribution of <i>so</i> per category.....	256
Figure 42: Distribution of <i>okay</i> per category.....	256
Figure 43: Frequency of <i>so</i> across NS and NNS in TaCoCASE	257
Figure 44: Frequency of <i>okay</i> across NS and NNS in TaCoCASE	258
Figure 45: Distribution of <i>so</i> as DM across individual speakers in TaCoCASE	259
Figure 46: Distribution of <i>okay</i> as DM across individual speakers in TaCoCASE	259
Figure 47: Distribution of functions of DM combination <i>okay so</i> in TaCoCASE	269
Figure 48: Distribution of new functions of DM combination <i>okay so</i> in TaCoCASE.....	271
Figure 49: Cultural aspects in TaCoCASE	284
Figure 50: Example topics from conversations about cultural aspects.....	290
Figure 51: Strategies for mediating cultural concepts	310
Figure 52: Key concepts operationalized for linguistic competence (CEFR 2020: 131).....	345
Figure 53: Results for collocates of <i>really</i> in WebCorp	348
Figure 54: Key concepts operationalized in the CEFR for discourse competence (CEFR 2020:138)	352
Figure 55: Key concepts operationalized in the CEFR for interaction strategies (CEFR 2020: 87)	358
Figure 56: Key concepts operationalized in the CEFR for explaining a new concept (CEFR 2020: 118).....	366
Figure 57: Criteria for building pluricultural repertoire (CEFR 2020: 124)	377
Figure 58: Successful online communication (CEFR 2020: 84)	389

Figure 59: Non-verbal communication in the CEFR (CEFR 2001: 89)	394
Figure 60: Worksheet vocab profile - task	404
Figure 61: Worksheet vocab profile	405
Figure 62: Worksheet vocab profile (teacher version)	406
Figure 63: Worksheet I - Explaining a new concept - task.....	411
Figure 64: Worksheet II - Explaining a new concept – strategies.....	412
Figure 65: Worksheet III - Explaining a new concept (student version).....	413
Figure 66: Worksheet let's talk about... customs and traditions	414
Figure 67: Worksheet Let's talk about ... TV series	415

LIST OF TABLES

Table 1: Operator verbs	16
Table 2: Descriptors for spoken interaction	37
Table 3: Communicative language strategies in the CEFR (CEFR 2020: 35)	53
Table 4: Cultural dimensions	58
Table 5: Descriptors for building up pluricultural repertoire	78
Table 6: ICC in federal state curricula	89
Table 7: Intercultural speaker and ICC objectives (Byram 1997:57; 2021: 84)	102
Table 8: ViMELF (Diemer et al. 2018)	155
Table 9: TaCoCASE (Collet 2023)	164
Table 10: Comparison of ViMELF and TaCoCASE	165
Table 11: Conversation topics	166
Table 12: TaCoCASE conversations	167
Table 13: Personal information of the participants in TaCoCASE	171
Table 14: Recording phases	183
Table 15: Transcription phases	184
Table 16: Examples of general CASE transcription conventions	187
Table 17: Examples of specific CASE transcription conventions I	188
Table 18: Examples of specific CASE transcription conventions II	189
Table 19: Benefits of TaCoCASE	194
Table 20: Frequency list TaCoCASE	219
Table 21: TaCoCASE and ViMELF - Word frequency lists	221
Table 22: Comparison of vocab profiles for TaCoCASE and ViMELF	229
Table 23: Vocab profile TaCoCASE	230
Table 24: Vocab profile ViMELF	230
Table 25: Common lexical bundles in conversation	241
Table 26: Collocates of <i>really</i> in TaCoCASE	244
Table 27: 3-grams in TaCoCASE	246
Table 28: 4-grams in TaCoCASE	246

Table 29: Concordances of <i>I don't know</i>	249
Table 30: Documented functions of <i>so</i> in NS and NNS contexts.....	253
Table 31: Frequency of <i>so</i> and <i>okay</i> in TaCoCASE.....	255
Table 32: Frequency of <i>so</i> and <i>okay</i> in ViMELF	255
Table 33: Collocates of <i>so</i> in TaCoCASE.....	260
Table 34: Frequency of DM combination <i>okay so</i> in TaCoCASE and ViMELF	267
Table 35: Concordances for DM combination <i>okay so</i>	268
Table 36: Functions of DM combination <i>okay so</i> in TaCoCASE	269
Table 37: Functions of DM combination <i>okay so</i> in TaCoCASE across and within turns.....	270
Table 38: Use of DM combination <i>okay so</i> across NS and NNS (same speaker)	273
Table 39: Cultural aspects in the conversations.....	288
Table 40: Further aspects of customs and traditions in TaCoCASE	299
Table 41: TV shows, series and movies in TaCoCASE	304
Table 42: Discourse structure of explaining (new) concepts	322
Table 43: Screenshots from WebCorp - comparisons.....	348
Table 44: Selected scales and respective DM functions.....	353
Table 45: Benefits of TaCoCASE in a teaching context.....	416

LIST OF ABBREVIATIONS

AFL	Academic Formulas List
AI	Artificial Intelligence
AWL	Academic Word List
BELF	Business English as a Lingua Franca
BI	Participant from Birmingham City University, Birmingham (United-Kingdom)
BNC	British National Corpus
BO	Participant from Boise State University, Idaho (United States of America)
CA	Conversation Analysis
CANCODE	The Cambridge and Nottingham Corpus of Discourse in English
CASE project	Corpus of Academic spoken English project
CBP	Corpus-Based Pedagogy
CEFR	Common European Framework of Reference for Languages
CL	Corpus Linguistics
CMC	Computer-mediated communication
COCA	Corpus of Contemporary American English
COE	Council of Europe
CoRD	Corpus Resource Database
DA	Discourse Analysis
DDL	Data-driven learning
DESI	Deutsch Englisch Schülerleistungen International
DigiCompEdu	Digital Competence Framework for Educators
DM	Discourse markers
DMIS	Development Model of Intercultural Sensitivity
EFL	English as a Foreign Language
ELF	English as a Lingua Franca
ELFA	Corpus of English as a Lingua Franca in Academic Settings
ELT	English Language Teaching
ENL	English as a Native Language
ESL	English as a Second Language
ESP	English for specific purposes
FL	Participant from University of Bologna, Forlì (Italy)
FLE	Foreign Language Education
FREPA	Framework of Reference for Pluralistic Approaches to Languages and Cultures
HE	Participant from University of Helsinki (Finland)
ICAME	International Archive of Modern and Medieval English
ICC	Intercultural communicative competence
ICE	International Corpus of English
KMK	Kultusministerkonferenz (Conference of education ministers in Germany)
KSAO model	Knowledge, Skills, Abilities and Other characteristics model
KWIC	Keyword In Context
L1	Native language of a speaker
L2	Second or Foreign language of a speaker
LL	Log-Likelihood
LOB	London-Oslo-Bergen corpus
MICASE	Michigan Corpus of Academic Spoken English
NNS	Non-native speaker
NP	Noun phrase
NS	Native speaker

NVE	Non-verbal element
OECD	Organisation for Economic Co-operation and Development
OER	Open educational resources
PISA	Programme for International Student Assessment
pmw	per million words
POS	Part of speech
RQ	Research Question
SAMR model	Substitution, Augmentation, Modification, Redefinition model
SB	Participant from Saarland University, Saarbrücken (Germany)
SF	Participant from St. Kliment Ohridski University, Sofia (Bulgaria)
SLA	Second Language Acquisition
ST	Participant from University of Santiago de Compostela (Spain)
STTR	Standard Type-Token-Ratio
SUE	Successful Users of English
TaCoCASE corpus	Transatlantic Component of the CASE project
TaLC	Teaching and Language Corpora (conference)
TBLL	Task-based Language Learning
TEFL	Teaching English as a Foreign Language
TESOL	Teaching English to speakers of Other Languages
TPACK model	Technological Pedagogical and Content Knowledge model
TTR	Type-Token-Ratio
TV	Television
UCREL	University Centre for Computer Corpus Research on Language
UK	United Kingdom
U.S.	United States (of America)
USA	United States of America
V	Verb
ViMELF	(Corpus of) Video-Mediated English as a Lingua Franca Conversations
WebCorp LSE	WebCorp Linguist's Search Engine
XML	Extensible Markup Language

1. INTRODUCTION

The world around us changes and teaching and the school system does not always keep up with these rapid challenges. Globalization, migration and digitization are challenges that call for new competences and it is the education system's duty to prepare learners in the best possible way for their future. Learners need to be educated so that they are able to live (and work) responsibly in a globalized and digitalized world (KMK 2012: 11). Teaching and in particular English as a foreign language teaching (EFL) needs to face these challenges and keep up with them – EFL teaching needs an upgrade.

There are several indicators that call for an upgrade. Textbooks and teaching materials usually do not provide authentic language examples, which leads to a gap between the language that is used in textbooks and authentic English that is used in conversations or in the real world (see for example O'Keeffe et al. 2007, Römer 2011, Mukherjee 2002). It seems that the "language presented in textbooks is frequently still based on intuitions about how we use language, rather than actual evidence of use" (O'Keeffe et al 2007: 21). Learners need to learn authentic phrases and expressions in order to communicate in an authentic way in situations outside the classroom (see Römer 2011; O'Keeffe 2007; Hunston 2002). Communicative competence is still one of the most important competences in the EFL classroom (Byram 2021). Traditional methods of teaching vocabulary and grammar are challenged when new digital tools, such as simultaneous translation of online meetings and AI tools that write summaries, are on the market. The use of these tools has an impact on the way we teach as well as on the way we assess competences. Intercultural competence is on the rise in a globalized world and may become – at least as important or – even more important than communicative competence in the future. In teaching, the focus is still predominantly on establishing communicative competence and only to a minimum extent on teaching intercultural competence (see Byram 2021). Thus, this imbalance needs to be countered and intercultural competence needs to be integrated even more, using authentic examples in order to depict real world intercultural

communication. In the era of digital transformation, digital competences play an important role in our lives and these need to be established in every subject – also in the EFL classroom (KMK 2016).

The central question that this dissertation tries to answer is the following: How can EFL teaching and learning be modernized and upgraded? One possible way of modernizing EFL teaching is suggested in this work: Using a corpus of video-mediated communication to improve communicative and intercultural competence in EFL teaching. My hypothesis is that using corpora and data-driven learning offers a powerful way to modernize EFL teaching. Corpora are collections of texts that depict authentic language in use. These texts – written and spoken samples – are stored electronically and can be analyzed with the help of computer programs (Hunston 2002; O’Keeffe et al. 2007; McEnery & Hardie 2012). Using corpora in a classroom and having learners investigate language in the corpus with the help of computer tools is called data-driven learning (DDL) (see Aston 2000; Mukherjee 2002; Gilquin & Granger 2010; see section 2.2.2.3.).

I argue that, depending on the text sample, corpora can help establish the main competences in EFL teaching – communicative and intercultural competence and to some extent also digital competences – all at once. I will therefore investigate to what extent the use of corpora in EFL teaching – corpus data and data-driven learning – contributes to improving communicative and intercultural competence and how a corpus of authentic spoken computer-mediated communication (CMC) – TaCoCASE – can be used in the EFL classroom. Examples from the corpus will be used to demonstrate the syllabus fit of it and to show possible applications of TaCoCASE in the EFL classroom in a very hands-on approach. The aim of this work is threefold: (1) providing new material that can be used in the EFL classroom, (2) suggesting and applying new methods through the use of corpora and DDL and (3) promoting a new mindset for a modern way of EFL teaching.

New material, new methods and a new mindset

The **new material** that is suggested here are corpora in general and in particular the corpus, TaCoCASE (Collet 2023), that has been specially designed for the purpose of this work. TaCoCASE is a spoken CMC corpus of intercultural encounters between students from Great Britain, Germany and the United States that were held via Skype. The corpus is a rich data source and, since it is available online, researchers can use the corpus and apply it in different fields of research. I will mainly use TaCoCASE in the EFL teaching context, but the findings provide insightful results that could be used in other research or other fields of application as well. Throughout this dissertation, I will present excerpts from the corpus that serve as examples for teaching aspects of communicative or intercultural competence. These examples depict authentic language in use and authentic intercultural situations in which the participants negotiate meaning – linguistically and culturally – and mediate (cultural) concepts. Additionally, the CMC conversations and their multimodal set-up sheds light on new aspects of real-life communication, such as non-verbal elements or strategies in online interaction, that are worth teaching (Byram 2021). The examples could be used directly in the EFL classroom and they could be turned into teaching material to include new aspects of, for example, digital or CMC communication and to enhance authenticity not only in the classroom but also in textbooks.

Suggesting and applying **new methods** includes (1) using corpus approaches, corpus examples and evidence-based findings from the corpus in EFL teaching (see for example McEnery & Hardie 2012; O’Keeffe et al. 2007; Hunston 2002) and (2) promoting data-driven learning as a new and learner-centered approach in teaching (Aston 2000, Johns 2002; Mukherjee 2002, Gilquin & Granger 2010). Even though DDL had been around quite some time, it is still not used in Germany very much in the EFL classroom or in teacher education. Maximizing the use of DDL and making it more accessible is as important as highlighting its effectiveness in EFL teaching. Using corpus methods means looking at frequency of words and statistical information about how words are used in combination, which is important information that should play a key role in selecting teaching material (Biber & Reppen 2002). These methods provide evidence-based findings about actual language in use that could either inform

textbook designers or teachers about what words are frequent – and therefore relevant – that should be included in textbooks or examples in the classroom (Römer 2011). Corpus methods and analyses thus come up with relevant words and expressions that should be taught in the EFL classroom. The conversations in TaCoCASE depict authentic intercultural encounters and excerpts from these conversations show how the participants negotiated meaning and exchanged cultural concepts.

Using corpora in general and TaCoCASE in particular has many benefits in EFL teaching, but one of the main benefits is that it facilitates data-driven learning. DDL is a method using corpora in the classroom in an action-oriented and learner-centered way so that learners become researchers themselves (Aston 2000; Mukherjee 2002; Gilquin & Granger 2010). It draws on communicative as well as digital competences at the same time as learners are working with corpora to investigate actual language in use and discuss their findings in class. It brings authenticity to the classroom through real-life examples, which textbooks or invented examples usually do not contain and learners investigate real language data themselves (Johns 2002; Gilquin & Granger 2010). They formulate hypothesis about language, test them and come up with findings and suggestions for their own language use, which is a very action-oriented and learner-centered way of learning languages (see Mukherjee 2002; Gilquin & Granger 2010) that can be very motivating for learners (Deci & Ryan 1993). Excerpts from TaCoCASE can serve as examples in the EFL classroom to establish communicative competence, intercultural competence or digital competence when completing a DDL activity. This means that all important competences are applied at once, which is very unusual and a major benefit for EFL teaching.

In general, the combination of using corpus material, corpus examples and corpus methods in the EFL promotes a **mindset change** in the EFL classroom. It brings authenticity to the classroom as material and examples are based on actual language in use (Gilquin & Granger 2010). Learners are exposed to authentic language in use and examples from corpora, which is usually not the case in textbooks. Using corpus methods means dealing with evidence-based information about words and expressions, which makes the selection process of words and phrases more authentic and guarantees a high relevance. DDL activities make learners researchers in a discovery learning and task-based language learning

(TBLL) approach that not only improves communicative competence, but also digital competence as well as future skills – communication, collaboration, creativity and critical thinking. This multi-faceted way of drawing on several major competences at once is one of the major benefits of including DDL in language learning. The mindset and the idea that underlies this work is that authentic input – be it linguistic or (inter)cultural – facilitates authentic output (Krashen & Terrell 1988; Swain & Lapkin 1995; Long 1996). Producing authentic output, not only in a communicative sense, but also in the sense of behaving in an authentic way in intercultural situations and being able to use digital tools in an authentic way, is the ultimate goal in language learning (see KMK 2012; KMK 2016). Establishing intercultural competence is very important even though the focus in EFL teaching is still predominantly on communicative competence. Putting more emphasis on establishing intercultural competence and adopting a discursive approach to intercultural learning (see Hallet 2008; König et al. 2022) is one major claim that I make in this dissertation. Consequently, I argue that there needs to be a shift from the native speaker as a model in language learning to the intercultural speaker (see Byram 2021), as the latter model combines communicative and intercultural competence (see section 2.1.5.2.7.). The results of my analysis indicate that the difference in language use between native speakers (NS) and non-native speakers (NNS) are not very large and that the special intercultural setting in TaCoCASE is one possible reason for that, which underpins the fact that the participants should be regarded as intercultural speakers rather than native and non-native speakers.

In this dissertation I will explore how the use of new material, new methods and the new mindset contribute to a modern way of EFL teaching, drawing on traditional competences that are extended in light of the current challenges and in light of future skills, using new methods and a new mindset. With this dissertation I argue for applying a new and modern way of EFL teaching using corpora. The benefits of using corpora in the classroom have been outlined by various researchers over the past 20 years, but we still do not see many corpora in EFL teaching (see Jablonkai & Csomay 2023; Mukherjee 2004). This was also a major point of critique at the 2024 Teaching and Language Corpora conference (TaLC 2024), in particular in Tony McEnery's plenary. A lot of work needs to be done to convince educators

and policy-makers of the benefits that the researchers working with corpora in a teaching context are already convinced about. Teachers are willing to use corpora and they see the benefits of it in teaching, but they are reluctant to use corpora as they do not fit to textbooks or to the syllabus because the use of corpora is simply not included in syllabi (see Farr & Leńko-Szymańska 2023; Mukherjee 2004).

I try to bridge this gap by outlining how TaCoCASE fits to what the Common European Framework of References for Languages (CEFR) demands, which is the key source for language syllabi in the European Union. I look at different competences the CEFR suggests and I propose ways in which TaCoCASE can be used to improve these competences, with examples. Thus, I argue that corpora align extremely well with the competences outlined in the CEFR and that there is an (implicit) syllabus fit which should be acknowledged by policy-makers to make it an explicit demand to use corpora in teaching. In order to be even more convincing, I will provide ready-to-use material and I will suggest concrete lesson plans for the EFL classroom. Thus, for this dissertation, I have not only created new material and used new methods in my research, I will also prove that the new mindset, which I am suggesting, fits to what the CEFR and policy-makers prescribe. This is not only the rationale and motivation for my work, but it is my deepest conviction that we need this mindset change sooner rather than later. This line of argument is also the basic and underlying structure of this work: I will first introduce the new material, TaCoCASE, then analyze the corpus using corpus methods and apply the findings and examples from the corpus in the classroom in order to show various direct applications and to prove the syllabus fit.

As far as my **methodology** is concerned, I am following an interdisciplinary approach, combining corpus linguistics, discourse analysis, pragmatics and conversation analysis with applied linguistics and EFL teaching methodologies. I will first analyze the language in the corpus and how the participants mediated cultural concepts, before applying these findings and the corpus itself in an EFL teaching context. In the analysis, I will use quantitative methods as a starting point in the analysis and conduct basic corpus linguistic analyses, such as frequency analysis, collocation and n-gram analysis as well as concordance analysis in order to identify frequent as well as relevant words and patterns. I then add a qualitative analysis in order to reach more nuanced findings. In this qualitative analysis, I will look at

how these words and patterns were used and what role and functions they have in a communicative way – on the discourse level – as well as in an intercultural way – when negotiating meaning, building rapport and creating common ground. In the application part, the focus lies on adopting the evidence-based findings from the analysis in the EFL classroom and providing examples from TaCoCASE that could be used in teaching. I will illustrate how examples from TaCoCASE meet the CEFR guidelines in terms of communicative and intercultural competence and thus prove the syllabus fit of TaCoCASE.

The detailed **structure** of the work is as follows: There are three main chapters, that build the core, and two additional chapters – introduction (chapter 1) and conclusion (chapter 6) – that set the work into a wider context. Chapter 1, this introduction chapter, describes challenges in EFL teaching and how this work is designed to solve these problems. The path towards a modern way of EFL teaching using corpora in the EFL classroom and focusing on communicative, intercultural and digital competences at once is outlined and the aim and rationale of this work is described as well as the structure and organization of this dissertation.

Chapter 2 is concerned with the theoretical background. It takes a look at research on EFL teaching and corpus linguistics as these two areas are combined in my interdisciplinary approach. The first major sub-chapter looks at aspects of (good) teaching and depicts the state of research in this field. The background and history of EFL teaching, main aspects of teaching methodology as well as commonly used frameworks and standards are described in order to understand the objectives of EFL teaching. Competences play an important role in EFL teaching and they are central elements in this dissertation. I will investigate how communicative and intercultural competences can be established with the help of corpora and corpus examples. These competences and their definition in the standards and frameworks will be described in great detail. EFL teaching in the digital era calls for new competences and in particular for digital competences, alongside the so-called future skills (Schleicher 2013; Poitzmann & Sobel 2024: 14). These aspects are crucial for modern and state-of-the-art teaching in the digital era. The second major topic in this chapter is corpus linguistics and its application in EFL teaching. The foundations of corpus linguistics, including corpus theory and corpus methods, will be

explored and I will look at how corpora can be used in the EFL classroom and what benefits and potential they have in EFL teaching.

Chapter 3 focuses on the **data and the methodology** used in this study. In the first part of this chapter, the corpus that is used and that was designed especially for this work – TaCoCASE – is introduced. The rationale, the design and the compilation process will be illustrated and some main corpus facts and figures will be presented. The methodology part of this chapter consists of a description and discussion of the methods that were used for the analysis.

In chapter 4, the corpus **TaCoCASE is analyzed** in terms of linguistic and (inter) cultural aspects. This chapter sets the basis for the application chapter and describes important aspects of language and culture that are depicted in the corpus. In using corpus methods and tools, the corpus is examined with regard to the language in use. Common corpus methods such as frequency analysis, collocation analyses and concordance analyses are conducted in order to find frequent words and patterns. In terms of pragmatic aspects of language, the use and functions of frequent discourse markers are explored. Differences in language use among native speakers and non-native speakers are also subject to analysis. As far as intercultural aspects are concerned, the focus lies on investigating certain cultural topics the participants talked about and the way they talked about them when negotiating meaning.

Chapter 5 illustrates **how TaCoCASE can be used in the EFL classroom** and to what extent communicative and intercultural – as well as digital competences – can be established when using examples from TaCoCASE. In the first part, examples from TaCoCASE are used to illustrate how the requirements in the Common European Framework of Reference (CEFR) are met in terms of communicative and intercultural competence. It is also explained how the special set-up of TaCoCASE sheds light on aspects that are usually not part of teaching, such as non-verbal communication and online interaction (Byram 2021). Suggestions how the examples can be integrated in the EFL classroom are described and two concrete topics are illustrated as case studies: how to teach vocabulary and

how to explain a (new) cultural concept. This chapter thus provides teaching ideas and material based on the findings in the analysis and based on the corpus TaCoCASE.

Chapter 6 – the **conclusion** chapter – recapitulates the findings of this work and puts them into a wider context. It summarizes the extent to which EFL teaching benefits from using corpora and in particular TaCoCASE and why it is crucial that EFL teaching is upgraded and modernized. Implications of this dissertation and an outlook for future research round off the chapter and the whole dissertation.

2. RESEARCH BACKGROUND

This chapter provides the **theoretical basis** and the **research background** for this work. It combines applied linguistics and aspects of teaching with linguistics and the field of corpus linguistics in particular.

Before looking at EFL teaching, some **general principles in didactics** will be explored. With a focus on the EFL classroom and EFL teaching, I will discuss what should be taught and what competences are important for the learners to acquire and establish. These norms are typically described in frameworks and standards; the two most important ones will be explored in greater detail, the Common European Framework of References for Languages (CEFR) on a European level and the German Education standards on a national level. Communicative and intercultural competences are – among other competences – very important in language teaching and also characteristic for the English as a Foreign language (EFL) classroom. I will investigate what competences are and what is meant by communicative competence and intercultural competence – starting out with general definitions and concepts and ending up at what the European and German frameworks and standards suggest for the EFL classroom. The EFL classroom is per se a room for learning how to cross linguistic and cultural borders and prepare students to live in a globalized world, to collaborate and communicate with people with different language and cultural backgrounds. Migration and globalization are issues in the EFL classroom and digitization and the digital transformation have in recent years also influenced EFL teaching. Therefore, I will explore how digital media and technology can be used in the EFL classroom and what needs to be done in order to integrate digital learning in language teaching.

A field that has had a major impact on the analysis of language and revolutionized the way language was analyzed is corpus linguistics. It looks at large samples of text of authentic language in use – corpora – and applies corpus methods in order to analyze language. The EFL classroom can benefit from this approach not only through the authenticity these texts bring with them, it can also draw the learners attention to corpus methods and analyses that offer new perspectives on language, which is

called data-driven learning. How corpora can be used in the EFL classroom and what influence they can have on language teaching are the two main aspects of this chapter. Using corpora and applying corpus methods in the EFL classroom thus means establishing new perspectives on language learning.

2.1. ASPECTS OF TEACHING

Teaching has many different aspects: What is the goal that should be achieved? What should be taught? How should the goal be achieved and how should content be taught? Questions of aims or objectives, content, methodology and technology are the guidelines in didactics and in teaching in general. This section looks at key terminology and positions ELF teaching with regard to its reference disciplines and its development as well as different levels of teaching and decision making processes. Some **objectives and competences** are prescribed by **frameworks and norms**. We will look at two major frameworks – the CEFR, on a European level and the education standards in Germany on a national level – and will look at what they define in terms of objectives and competences in EFL teaching. As far as competences are concerned, we will approach the term by looking at definitions from different fields ending up in the educational context and the definition used in Germany. Two major competences, which are key in the EFL classroom are communicative and intercultural competence. This section takes a closer look at the individual competences and the most important aspects of them. I will answer questions like what is meant by communicative and intercultural competence, what aspects belong to each of these competences and what aspects the CEFR and the education standards list that belong to these competences. Whenever possible, I will refer to currently ongoing debates with regard to goals in language teaching – for example what speaker model should be used – and I will discuss to what extent the prescribed norms are feasible or sufficient or should be expanded in the future. The last sub-chapter focuses on EFL teaching in the digital era. Digitization and digital education affect all subjects and guidelines for EFL teaching have only been recently established. What is actually meant by digital transformation, what aspects influence digital transformation and

some key terms will be described. Consequences and challenges of the digital transformation for EFL teaching will be discussed and suggestions for teaching will be made.

2.1.1. DIDACTICS AND DIMENSIONS

Teaching is often referred to as **didactics**, which stems from the Greek words “didaskalos”, meaning teacher, and “didaktiké téchne”, which translates as ‘the technique of teaching’ (Decke-Cornill et al. 2015: 2). This definition puts the teacher in the center of the whole process and looks at teaching in a holistic and creative yet appreciative way. Another widely used expression is ‘teaching and learning’, as it also sheds light on learners as well and implies that learning is not only what the teacher says but also how learners profit from the whole process. Decke-Cornill et al. (2015) emphasize the institutional perspective in their definition of didactics as “official, planned, coordinated and compulsory teaching and learning at school” (Decke-Cornill et al. 2015: 2). This puts didactics into a school context and also highlights that what is being taught follows a certain structure or a plan, and is very carefully chosen by teachers who are responsible for this whole process of teaching and learning. Didactics can generally be any kind of teaching and learning and is not restricted to institutions or fields – including further education – even though the most common context is school. With the rise of lifelong learning, continuing education in companies and teaching and learning at universities, the term didactics can also be applied to those fields (e.g. *Hochschuldidaktik*). The basic principles of didactics are the same for each applied context. Cohn and Terfurth (1997) came up with a **didactic triangle** between teachers, students and the content that is taught at school (cited in Decke-Cornill et al. 2015: 2). From this triangle, didactics derives three main tasks: (1) using content in order to offer learning opportunities and activities, (2) teaching and the professionalization of teachers through further education, and (3) learning, which means that with a focus on the learner, learning conditions should be improved and the learners’ situation needs to be taken into account so that the learning process can fully unfold.

Closely related to this didactic triangle but focusing on slightly different aspects, are the three **dimensions of didactics** as Decke-Cornill et al. (2015: 4) stated them, namely

- the *analytical dimension* (What happens in class / in teaching?)
- the *normative dimension* (What should be taught? What should happen in teaching?)
- and the *operative dimension* (methodology and methods, i.e. how educational processes can be encouraged).

Looking at the triangle and the dimensions, the intersection of these models results in the three main aspects of didactics: objective, content and method (Haß 2006: 69), which are carried out, selected or set by the various people involved in the educational process (teachers, learners and governments).

Objectives, Content and Method

After having looked at various definitions of didactics, it is now important to dive a little bit deeper and see what the core of didactics actually is and means. Before doing that, one terminological aspect needs to be discussed. Sometimes, didactics is also referred to as 'methodology' but the term 'methodology' should be carefully differentiated and not mistaken with the term 'methods' as the latter is only one aspect of didactics. The three main aspects, **objective**, **content** and **method** (Haß 2006: 69) will be explained in further detail here.

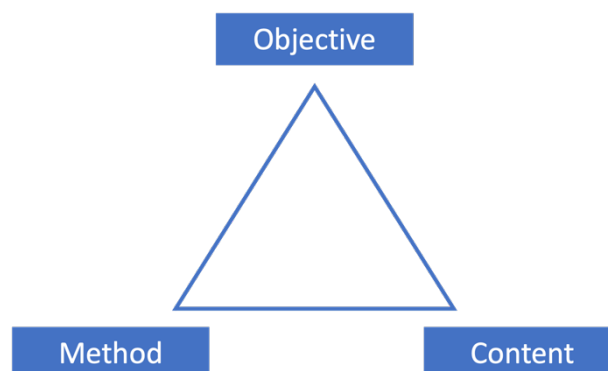


FIGURE 1: DIDACTIC TRIANGLE

Objective

Like in every company or in life in general, people need **objectives** that they can work towards and people who set objectives, who need to check whether the objectives were reached or not. If the objectives were reached, quality is ensured and the development (in the company or in life) can continue and new objectives can be defined. If the objectives were not reached, an analysis or evaluation is needed in order to determine a strategy to reach the objective in a second round.

With the rise of evaluation and assessments in schools, it is important to set objectives to be able to compare learner output across different classes in one school, different schools in one federal state as well as different federal states or even different countries (see for example student assessment like in PISA). Objectives on a macro-level (state, country-level) are usually defined in terms of competences. Objectives on a micro-level usually concern the school or classroom level. While competences are political guidelines defined by the European commission (Common European Framework of Reference) or the German education ministers (Educational standards issued by the KMK) (see sections 2.1.4.1. and 2.1.4.2.), objectives in the classroom are usually defined by teachers themselves. Depending on the curriculum, there are several objectives teachers should define: objectives which students should reach at the end of a school year, at the end of a unit (whatever is defined as a unit) or at the end of a lesson. These objectives correlate with and built up on each other so that students reach the requested competences (see section 2.1.5.) after a defined period of time.

Objectives are usually defined by teachers and fulfill two main criteria: structure and assessment. As far as structure is concerned, objectives help the teacher to structure their units and plan their lessons, for example how many lessons they need in one unit in order to reach a certain objective. They also help learners to understand what is expected of them and makes the whole learning process more transparent. It is easier for teachers to justify in exams and assessment tests, whether learners did or did not reach certain objectives. This is also important for learners to understand where they have deficiencies and where they need to put some more work into certain topics.

Setting educational objectives

The most widely used taxonomy for setting **educational objectives** was established by Benjamin Bloom (1956) and consists of six levels with different cognitive domains:

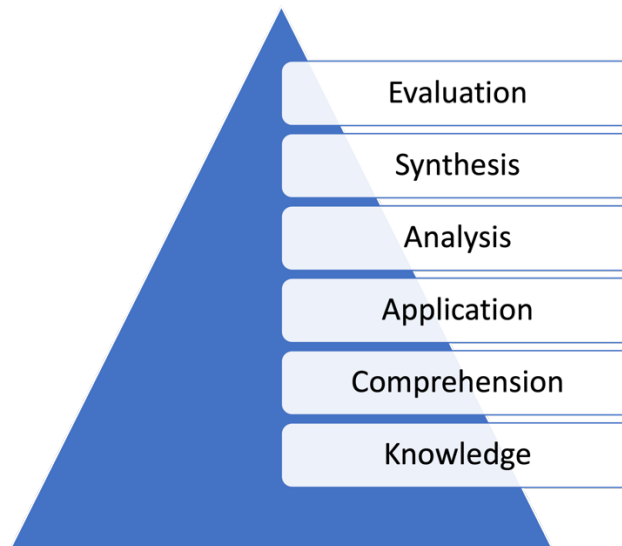


FIGURE 2: EDUCATIONAL OBJECTIVES (BLOOM 1956)

The higher the domain (see figure 2), the higher the cognitive involvement and thus the learning outcome for learners. When defining objectives, it is important to address different cognitive levels and not only lower levels as the learning outcome is better if there is a mix of different cognitive domains. When formulating objectives, the “can-do-approach” should be used as well as active verbs, or the so-called operator verbs.

Operator verbs are, on the one hand, used to describe the objective and, on the other hand, used when formulating assessment questions. This taxonomy is also often used as basis for assessment and especially in (oral or written) exams to vary levels of difficulty in the exam tasks. Table 1 gives an overview of different operator verbs for each level according to Bloom’s taxonomy (see Meyer 2004):

TABLE 1: OPERATOR VERBS

Level		Operator verbs
1	Knowledge	Name, tell, write down, ...
2	Comprehension	Interpret, explain, translate, identify, define, present, ...
3	Application	Apply, find out, plan, calculate, solve, ...
4	Analysis	Isolate, determine, analyze, test, compare, ...
5	Synthesis	Connect, combine, develop, design, ...
6	Evaluation	Decide, argue, classify, evaluate, ...

- **Example for a lesson objective :** The students know at least five sights in London. The students can name at least five sights in London.
- **Example for a task in an exam:** Name three sights in London.

Typically, there is not only one lesson (or unit) objective, but there are more objectives from different cognitive levels. These cognitive levels are also drawn upon by curricula in Germany, where usually three different “**Anforderungsniveaus**” or **levels of requirement** are defined: a minimum level, a regular level and an expert level. On the minimum level, reproducing information is key, which means that learners can apply basic knowledge and can repeat information they learnt before. This is very much on the level of declarative knowledge and draws on the categories “knowledge” and “comprehension” in Bloom’s model. The regular level includes reconstructing knowledge. This includes linking knowledge and drawing connections. The regular standard comprises levels three and four in Bloom’s model including application and analysis as educational objectives. Expert status or the highest level of requirement has to do with transfer. Strategies for solutions are being developed and interpretation and evaluation is included. The expert level corresponds to level five and six in Bloom’s model and consists of synthesis and evaluation.

These levels of requirements also need to be assessed in order to evaluate the learners' input and output. Assessment will be discussed later in this chapter as part of the CEFR and proficiency levels. However, it is important to mention already that setting objectives and evaluating or assessing if the objectives were attained or not is important. This is why these objectives serve for the teacher as a starting point for teachers when planning lessons, but also as a guide throughout the learning process and in the assessment. Typically, reproductive exercises or exam questions only ensure a minimum level of requirement and thus correspond to grade four in the German grading system. Expert level draws on transfer questions in the exam and only if this question is answered correctly, the learner achieves the complete learning outcome and thus earns the highest grade, meaning a "one" or "very good" in the German grading system. Again, this helps the teacher and the learner to evaluate at which level in the learning process the learner is at and how much guidance is needed in order to reach the full outcome.

Content

The second dimension describes the **content** that should be learnt or taught. Typically, the main content can be found in the respective curricula, for example "challenges of a global world" would be one topic which is addressed in the curriculum (KMK 2016). What belongs to this topic is usually quite open for the teachers and they can decide what they would like to cover and what they think is useful to know in order to fully understand a topic. Climate change, for example, would be a piece of content which could be covered and with it vocabulary, background information and other facts like temperature statistics, the effect of rising temperature on the ecosystem and so on. This content can be presented through texts or videos about climate change and the decision on how to present content and what material to use is the teacher's decision. When choosing content, it is important to have a look at the curriculum and also to think about what the target group or the learners need to know. According to Haß, content should be clear, the topics should be interesting for the students as well as adapted to their respective age (Haß 2006: 69).

For example, in order to take part in a discussion on global challenges, the teacher could decide to talk about climate change in particular. The students need to acquire specific vocabulary, they need to know what climate change is and what challenges exist, and make up (or search for) solutions people can adopt in order to address the threats. Facts as well as vocabulary belong to the content in this case and are probably of interest for the students as it is covered in the media and it affects their behavior and their future.

Method

The term method stems from the Latin/Greek word ‘methodus / méthodos’ meaning the way to achieve a goal or as Doff puts it “a path towards a goal” (Doff 2018: 4). It is important to notice that the term methodology is different from method. Methodology is a synonym for didactics, meaning teaching and learning in a general sense. However, method is only one aspect or part of methodology. “[A] method answers the question of how teaching and learning are organised” (Doff 2018: 4) and thus describes the way of teaching and learning leaving the two other factors – objectives and content – aside. According to Haß, using various methods is a sign of good teaching and speaks for quality in lessons (Haß 2006: 72).

Methods thus include decisions on how content is presented and how active the students should be in the learning process. This is first of all a question of learning arrangement (“Lernform” or “Sozialform”). The teacher could for example explain or present a certain topic and the students listen to him/her (“Frontalunterricht”). The learners could work together in groups (cooperative learning) or they could work individually and read a text or watch a video (individual learning). These methods could also be combined so that a mixture between “Frontalunterricht” and cooperative learning in group work is applied in a lesson.

As far as **cooperative learning** is concerned, there are several different learning activities and learning methods (“Lernmethoden”) which could be carried out, for example “group puzzle” (sometimes also referred to as jigsaw puzzle), “bus stop” or “double circle” (see Haß 2009), which are different kinds of

collaboration. Cooperative learning usually follows the principle: Think – pair – share (Haß 2009: 4). This means that the learners start with a task they carry out individually, like for example reading a text with information about a specific topic. In the second step, “pair”, they form groups of two and discuss their findings. In a third step, they share the results with their classmates so that everyone has the same information and everyone presents “their” topic to the class. There are several characteristics that differentiate cooperative learning from group work: positive interdependence, face-to-face interaction, individual accountability, social skills and group processing (see Bonnet et al. 2017: 173; see also Johnson et al. 2008; Bonnet 2009). The main advantage of cooperative learning is that the students are active throughout the learning process and acquire knowledge on their own instead of passively listening to a teacher. They are more involved in the learning process themselves and also learn social competences when collaborating with other students or presenting information to others (Bonnet 2009). However, cooperative learning needs a lot of preparation beforehand by the teacher: They have to find texts that are of equal length for each of the groups, the teacher has to make sure that the students grasp all important information and present it in a way the other students can understand it. A discussion and a feedback phase are very important and should always be included. It is up to the teachers which learning method and which learning arrangement they choose, having in mind that it should fit to the content and the lesson plan.

To sum it up, the relationship between objectives, content and method is very close and they depend on each other. Objectives lead teachers and learners through the learning process where the acquisition of different competences is the ultimate goal. These competences are defined by policy-makers and prescribed in frameworks. Objectives – be it lesson or unit objectives – are defined by teachers and answer the question what the students should be able to do after a certain period of time. Another ‘what-question’ is answered by the content teachers provide and what the students should learn (or the teachers should teach) in order to reach the respective goal. The way or how to reach the objective belongs to the method the teachers can choose in order to deliver content (or let the students explore content) in order to reach a certain objective.

Technology

With the digital transformation in all areas of our lives, a fourth aspect to didactics, namely technology or digital tools could be added.

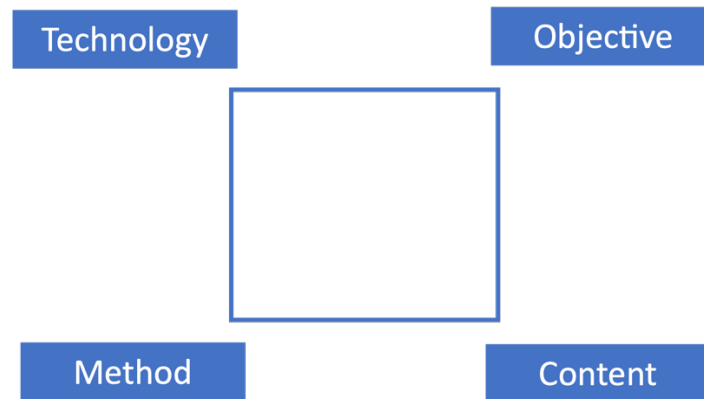


FIGURE 3: DIDACTIC SQUARE

One could argue that technology could be subsumed under method, but I would rather put it as a separate category as it is too important in terms of 21st century skills. Technology as it is understood here, comprises of digital tools, including digital media and tasks. In light of the 4K-model and 21st century skills (see section 2.1.6.), it is important to choose the correct digital medium and the respective technology so that technology and digitization support teaching. Neither the tool nor the technology is in the focus, but the learning process itself. Digital collaboration or communication tools can be integrated in the EFL classroom to establish general competences, communicative competence as well as digital competence. They can support the language learning process by offering new ways of presenting learning material and new tasks that activate and involve the learners even more in the learning process. Therefore, the decision about technology should be a central question in didactics¹.

¹ Integrating digital tools in the classroom will be discussed in greater detail in section 2.1.6.

2.1.2. FOREIGN LANGUAGE TEACHING

In teacher education, which is still the most common context for didactics, there exist didactics for each subject, such as chemistry didactics or maths didactics. In Germany, each one of them is called **“Fachdidaktik”** (translates to subject didactics or subject-specific teaching) and emphasizes the respective subject. Sometimes related subjects are combined like for example science or foreign languages, which is then called foreign language didactics (German: “Fremdsprachendidaktik”). Foreign language didactics or “foreign language education” (FLE) is the “academic discipline which deals with the teaching and learning of foreign/second language(s) (L2) in institutional, primarily school-based contexts” (Doff 2018: 10). These disciplines were established in Germany only after World War II (Surkamp et al. 2018: 10).

English, Spanish and French are very common foreign languages that are usually taught at school as individual subjects. Even though in some schools in Germany in grades five to seven of secondary education, science is taught as one subject including elements of biology, chemistry and physics, languages are usually taught individually for several reasons. There are some underlying principles in language education which are the same for every language and which are covered in foreign language didactics in general. However, most of the elements are unique to the respective languages, like vocabulary or grammar rules, which should be taught separately. Therefore, each language usually has its own didactics like for example **English language didactics**. The following figure (figure 4) displays the relationship between the different terms and areas of didactics.

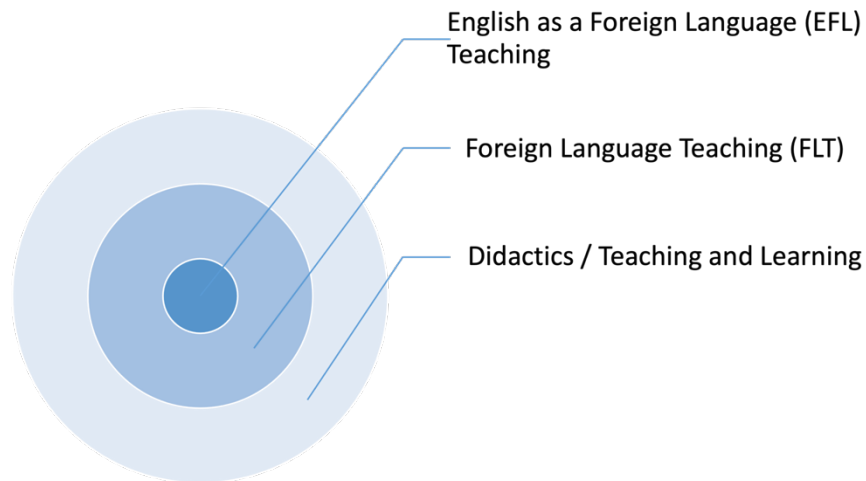


FIGURE 4: EFL TEACHING, FLT AND DIDACTICS

Didactics is the umbrella term and is concerned with general issues in teaching and learning, like objectives, content and methods. The way of teaching and the general underlying principles of teaching and learning, like how a lesson should be structured or how to motivate the students in order to make the most out of the learning process, counts for all subjects and is a skill every teacher should have. The closer we get to the core (see figure 4), the more specific becomes the teaching. It is the specific aims, content and methods for the respective subjects and the ‘fine-tuning’ that is important. It makes a difference if the teacher uses an experiment in chemistry to get the students’ attention or if the teacher shows a video of an intercultural encounter of two people in the English language classroom. The intention of motivating the students at the beginning of a lesson (see Deci & Ryan 1993) and using an inductive way is the same for both, but the content and the specific method is different and needs to be adapted according to the subject. One could also read figure 4 from the core to the outer circle, meaning that English language teaching belongs to Foreign language teaching and this again belongs to the wider field of didactics or teaching and learning.

2.1.3. ENGLISH AS A FOREIGN LANGUAGE TEACHING

In international contexts, English Language Teaching (ELT) or English as a Foreign Language (EFL) teaching are the most widely used term, while in Germany English language didactics (“Englische Fachdidaktik”) is most common. Other terms which are sometimes used are TEFL (Teaching English as a foreign language) or TESOL (Teaching English to speakers of other languages). All these terms have in common that the main subject of learning is English. The only difference is to what extent they give hints or highlight the target group, meaning the learners. I will use the term didactics for teaching and learning in general and English as a foreign language (EFL) teaching for what is referred to in Germany as English language didactics. If issues only count for schools or teaching in Germany, this will be highlighted then.

2.1.3.1. ENGLISH NATIVE AND (NON-NATIVE) SPEAKERS

It is important to take a step back and look at what “English” means. English is the most popular foreign language as it is the most widely spoken language on earth. According to Statista (2023), there are 1.5 billion people worldwide who speak English either natively or as a second language, followed by Chinese (Mandarin with 1.13 billion speakers worldwide), Hindi (609 millions) and Spanish (559 millions) speakers. And English is a world language, following David Crystal’s definition which says that a world language has a “special role that is recognized in every country” (Crystal 2003: 3). However, there exist different kinds of “English”, which were for example classified by **Kachru** (1985, 1992) in his model of World Englishes. He defined **three circles**: The inner circle, the outer circle and the expanding circle (see figure 5).

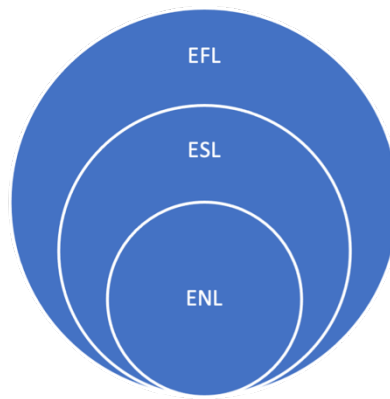


FIGURE 5: KACHRU'S CIRCLES OF ENGLISH

The inner circle represents **English as a native language (ENL)** and thus focuses on native speakers of English, which means that they stem from the “traditional bases” where English is the “primary” language (Crystal 2003: 60). Examples would be countries like the United States, the United Kingdom, Ireland or Anglophone Canada. A native language or L1 (first language) in general is the language a speaker is born with or “[t]he first language a human being learns to speak” (Bloomfield 1993: 43; cited in Cook 2008: 171). Thus, a **native speaker (NS)** – of any language, not only English – is a person who speaks a language “since early childhood” (McArthur 1992, cited in Cook 2008: 171). The outer circle, sometimes also referred to as “World Englishes”, represents **English as a second language (ESL)** and originates from countries, which were former colonies and where English is still the official language, even though inhabitants of these countries also speak other native languages. ESL is often used for intranational communication (i.e. communication within one nation) as a lingua franca (Meierkord 2012) as well as international communication (communication between speakers of different nations). There are different **varieties of English** in the **Outer circle** and their development has been described by Schneider’s dynamic model (Schneider 2011). In this model, a sociolinguistic perspective is undertaken and language use is an act of identity. Non-standard or new forms in some varieties of English in the Outer circle include differences in adverb position, postponed conjunctions, numbers, irregular plural forms and singular forms, e.g. **furnitures* and **trouser*) and reduplication as emphasis (Crystal 2003: 153). Examples for ESL countries would be India, Nigeria or Pakistan. In ESL, new norms

and standards in contrast to ENL are developed (Hundt & Mukherjee 2011: 2; Schneider 2011), for example different phonetic or lexical features. In these cases, English is used as a **second language**. People acquired English besides their native language and use it in their country for administrative purposes, at work or at school. Using English as official language at school should not be confused with learning English at school, where English would be a foreign language.

The third circle, the expanding circle (see figure 5), represents **English as a foreign language (EFL)** and is usually represented by learners of English, i.e. English is not their native language and it does not have an official status in their country even though it has a special role as it is used as a means of communication in different contexts like business or in some subjects at school. Examples would be countries in Europe except the Anglophone countries, China, Russia or Japan. Learners of a specific language are usually described as **non-native speakers (NNS)** of that respective language, as they are learning a language in addition to their native language. Due to globalization, the number of speakers in the expanding circle is still rising as more and more people are learning English.

Typically, there is a distinction between L1 and L2 of a speaker. L1 is the native language of a speaker, as mentioned above. L2 is usually considered as second language or foreign language, which means that it is an additional language (or more than one additional languages) a speaker knows other than their native language.

Having defined the categories of English, the question now is where ELF comes in and what kind of “English” we should aim at when it comes to teaching. This is also important in terms of classification of the speakers of the data in the corpus used in this dissertation.

In many business contexts and also in private context – when people with different language backgrounds meet – English is used as a medium to communicate or as lingua franca. **English as a Lingua Franca (ELF)** was first regarded as a “variety of English” (Jenkins 2015b: 77), which would place it in the category of World Englishes and thus in the outer and expanding circle in Kachru’s model. Barbara Seidlhofer and Jennifer Jenkins have been working on a new understanding of ELF, focusing

on the transcendence of English as it crosses (national) borders and the negotiation of meaning between interlocutors (Jenkins 2015b: 73). Seidlhofer defined ELF as the “preferred option for cross-cultural communication” (Seidlhofer 2003: 9) and she opens up the term by stating that ELF is “any use of English among speakers of different first languages for whom English is the communicative medium of choice [...]” (Seidlhofer 2011: 7). Thus, Seidlhofer and Jenkins leave national boundaries and the geographical backgrounds of speakers aside and rather look at communication between speakers with different language backgrounds and intercultural encounters. The central aim in ELF is having a set of strategies to “achieving mutual comprehension” (Mauranen 2012: 7) and successful communication (Jenkins 2015a: 45) between speakers of different language and cultural backgrounds. Successful communication also means that there could be non-standard features and that “mistakes” or non-standard forms are rather seen as “evidence of linguistic adaptability and creativity” (Jenkins 2015a: 45) of ELF speakers and not as deficiencies. ELF is characterized by a “high degree of interactional robustness, cooperation, consensus-seeking behavior and affiliation” (Firth 2009: 149), which makes sense in terms of negotiation of meaning.

When it comes to teaching English, the most important question is: **What kind of English should we teach at school?** What is considered as “standard” English? Should learners speak like native speakers? Or should learners be able to get a message across in English so that their interlocutor can understand them? These are probably the two ends of a continuum and during the development of English language teaching, the goals have changed and there is still a controversial debate on where the learners should be on that continuum.² ‘Near-native’ or ‘native-like’ competence was the goal for a long time, as this is also the characterization of C2 level in terms of language proficiency in the CEFR (see section 2.1.4.1.). Language proficiency needs to be assessed in some way and a standard against which learners can be compared needs to exist. This standard has been the native speaker for a

² See in this context section 2.1.5.2.7., which focuses on the native speaker as a model for language teaching. Here, in this section, the focus is on communicative competence only. Section 2.1.5.2.7 takes a broader view and discusses the intercultural speaker as a model, combining communicative and intercultural competence.

long time and I would argue that, in most cases and especially in younger years in school, ENL and British or American English in particular, is still the standard in Germany. However, new approaches in research discussions arise, which question whether the native speaker is still a valid reference (see for example Mukherjee 2005, Seidlhofer 2011, Shakouri & Shakouri 2014). Some popular arguments are that ELF speakers outnumber the native speakers of English (Graddol 1998), and that English is more significant in an international context than English as a first or second language, as learners usually use English as their foreign language to communicate. Another argument is that successful communication or use of English (see e.g. “SUE”, meaning successful users of English, coined by Prodromou 2005) is more important than correct grammar and spelling. Sometimes there is even a rivalry between ELF and EFL, but they are also intertwined to some extent. Mauranen makes this distinction very clear: students are EFL learners as long as they are in the classroom and they are ELF users as soon as they leave the classroom and become a user of English (Mauranen 2012: 5). This distinction between language user and language learner is key to the difference between ELF and EFL. Vivian Cook argues that “language teaching would benefit by paying attention to the L2 user rather than concentrating primarily on the native speaker” (Cook 1998: 185). I think this shift from the native speaker to the language user or learner has been very much implemented in the CEFR and their learner-oriented-approach (*learner-centeredness*) in formulating the competences. ELF-aware teaching is on the rise, as the number of studies around this topic increases (see Jenkins et al. 2018; Vettorel 2015). Raising awareness for the ELF mindset – ‘mistakes’ versus creativity – is very important and also leads to a more communicative way of teaching and learning English. However, there is still a lack of ELF in all phases of teacher education (see Lopriore & Vettorel 2015; Galloway 2018), i.e. there should be more ELF teaching material, more teacher training around ELF language and it should be part of the curriculum.

All in all, ELF and EFL both have rights in their own and both are needed when it comes to teaching. The context makes the difference. We need ELF teaching in the classroom as the students have to learn English and acquire the basic rules of English. Therefore, assessment at beginner level in school needs

a certain standard in terms of communicative competence and linguistic competence in particular. This standard will probably still be the native speaker, especially in form-focused phases like teaching grammar or spelling. However, we also need to prepare the students for the real world and for interactions in English. And in the real world, the students are *language users* in different contexts and they need to know which rules then count. These situations, when learners are using English outside the classroom, are either with native speakers or non-native speakers of English, which makes all these situations ELF situations (see definitions above). Students need communicative and intercultural strategies to perform in these real-world situations. As soon as other aspects of communicative competence (pragmatic competence, socio-linguistic competence) and intercultural competence are concerned, which is to a large extent at later stages in English language teaching, ELF language should be integrated. Students should therefore be familiarized with and sensibilized for ELF language and the ELF mindset. A good way of integrating ELF language and teaching communicative and intercultural strategies would be through exposing students to authentic language examples, which can be found in corpora like ViMELF and TaCoCASE. The goal of ELF teaching in the classroom could be to make the students a competent ELF user in the real world. What we need to teach in order to reach this goal and what competences a proficient English user should acquire, will be explored in section 2.1.5.

2.1.3.2. REFERENCE DISCIPLINES OF EFL TEACHING

English as a foreign language teaching is influenced by several **academic branches** like linguistics, cultural studies, literature and language education as well as didactics or pedagogy, including research in the fields of psychology (learning theory, language theory, brain research) and applied linguistics as well as second language acquisition (Haß 2006: 10) (see figure 6).

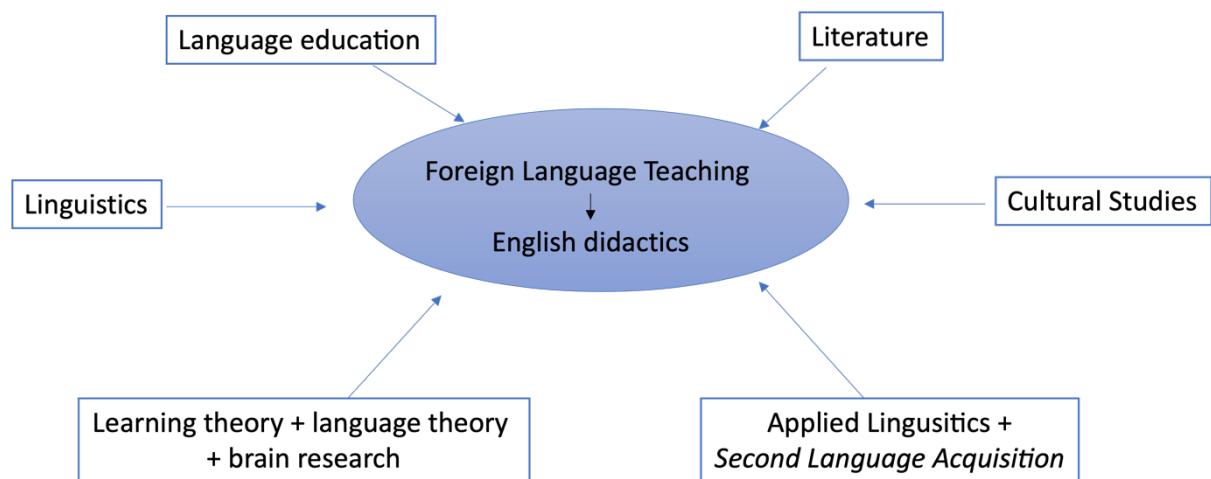


FIGURE 6: ENGLISH DIDACTICS AND ITS INFLUENCES (HAß 2006: 10)

The academic branches like literature, cultural studies, linguistics and to some extent also applied linguistics usually find their way into the curriculum through delivering content and findings on what should be taught. Language education, psychology with speech and learning theory and brain research as well as applied linguistics with second language acquisition are mostly methodological issues and concern the way of teaching and learning, thus didactics. Haß (2006) gives – very brief – insights into what these branches cover (notice that this list is not exhaustive but focuses on the issues covered in school contexts):

- Literary studies are concerned with literary theory and the study of literary texts for example poetry, drama, fictional and non-fictional texts. It also has a cultural aspect as the literary canon and literary texts give insights into different aspects of culture.
- Cultural studies are concerned with history, politics, anthropology and intercultural learning of specific cultures.
- Applied Linguistics is usually referred to as sub-branch of linguistics and is mainly concerned with English language teaching and (second) language acquisition (SLA), meaning the theories how (second) languages are acquired. There is another definition that describes applied linguistics in a more general way, meaning using insights from any branch of linguistics to tackle real world problems.
- Learning theory, speech theory and brain research focus are empirical fields and belong to the wider field of psychology. They are concerned with cognitive processes, how the learning process in the brain is working, what are motivational factors of learning and thus looks at the learner in the language acquisition process.
- Linguistics brings in elements of syntax, phonetics, morphology, pragmatics and socio-linguistics.

- Language education belongs to didactics and looks at teaching and learning languages at school, for example how languages are taught in the classroom with objectives, content and methods specific to languages.

These areas sometimes overlap and some elements from one branch also belong to another branch. For example the question of how we can motivate English language learners clearly belongs to language education but also draws on elements from the psychological branch. Second language acquisition can be subsumed under linguistics but is also influenced by speech theory or brain research.

2.1.3.3. DEVELOPMENT OF EFL TEACHING IN GERMANY

EFL teaching and subject-specific teaching (“Fachdidaktik”) in Germany only established after World War II (Surkamp et al. 2018: 12) when more and more researchers focused on teaching. It became an own discipline besides the subject disciplines and educational sciences. In Germany, many “Pädagogische Hochschulen” (translates to educational universities of applied sciences) emerged and focused on the education of teachers. English itself became a mandatory subject for students (cf. Hüllen 2005) and the need for English teachers was there. Since then, there have been constant changes with phenomena occurring and disappearing in EFL teaching. In the early years, the **focus was on form**, with behaviorism and structural linguists being influential around the 19060s. Grammar rules, patterns and habit formation were in the focus and the Grammar-Translation method, the direct method and the audio-lingual and audiovisual method with drill instruction especially in pronunciation were prominent methods during this time (Surkamp 2018: 19). All these methods focused on receptive and productive activities and understanding and repeating patterns was more important than speaking. In the late 1970s, structural linguists were under question and with this shift in mindset, a shift of focus towards a more communicative way of teaching English came along: **the communicative turn** (see Surkamp 2018: 19). This focus on communication and **communicative competence** was new

and pointed the way ahead as “[w]hat was needed in order to use language communicatively was communicative competence” (Richards 2006: 3). This new approach influenced EFL teaching as interaction played and still plays an important role in the classroom. Students should be able to communicate and be a competent language user. During the last years, with globalization and cultural diversity in our everyday life is rising, intercultural didactics and intercultural competence have become more and more important. EFL teaching has to some extent adapted to this trend already, even though much work needs to be done especially in terms of methodology and content.

It is important to mention in this context some general educational changes, namely the change from **focusing on input to focusing on output**. Focusing on input meant that syllabi and guidelines were installed with relative specific content regarding what should be taught. Focusing on output put the learner in the center of the learning process and competences the learner should acquire, are nowadays formulated in guidelines. This change can be attributed to increasing assessment and quality assurance in the education system. Nation and even Europe-wide comparison of learner performance was tested with studies like PISA or DESI. The goals were (1) to reach equality of opportunity across different European countries or federal states in Germany with the same criteria and common objectives and (2) to see where deficits in teaching or learning can be found in order to work more efficiently to reach certain objectives. Thus, comparison and evaluation are at the center of this new approach and frameworks and standards – on the European and national level – have been developed in order to reach these goals.

2.1.3.4. LEVELS OF TEACHING

Before looking at the frameworks and standards, it is important to look at the different **levels** where decisions about the learning process and about curricula are made. There are three levels of curricula

planning and decision-making processes: (1) **macro**, (2) **meso and micro** as well as (3) **nano** (COE 2016: 19).

- The **macro level** includes nations, states or regions. On the macro level, general aims of teaching as well as specific aims or competences are set. Also, decisions on content are made on this level, which entails decisions on what the learners are required to learn. This level is mostly concerned with the “what” question on learning. (COE 2016: 19)
- The **meso level** represents the school and the **micro level** the classroom level. In these levels, decisions about the “how” are made, which includes activities and approaches, grouping and location, time, resources, the role of teachers as well as co-operation. Decisions about how the students learn, with whom (cooperatively or individually), where, how much time they have in order to acquire the content and to what extent the learning process is guided by the teacher or by the learners themselves. (COE 2016: 19)
- The **nano level** represents the individual student and is mainly concerned with evaluation. This means that the learning process is evaluated and the degree to what extent the respective competences are acquired is assessed. (COE 2016: 19)

From these different levels of teaching arise different responsibilities. On the **macro level**, usually national or federal state governments are responsible for decisions on curricula and competences. The **meso and micro level** lies mainly in the responsibility of the teachers – in most cases the individual teacher, but it could also be the school principal administration. In charge for the assessment and evaluation on the **nano level** is the teacher, even though the learners of course have an influence on this level in terms of performance and how positive the evaluation is. It is important to have these levels in mind in order to address the right people in charge for different decisions concerning teaching and the design of curricula.

2.1.4. FRAMEWORKS AND STANDARDS

Teaching English in secondary education in Germany focuses on three main questions (on meso and micro level): What are the objectives the students should reach, what should be taught (content) and how it should be taught (methods and technology). Politics is usually there to set boundaries in education and to guarantee equal opportunities. This means that objectives are usually set at the political level – either in Europe or in the respective countries (macro level). The two main frameworks

which are important for the EFL classroom in Germany are the **Common European Framework of Reference** – on a European level – and in the **education standards** in Germany called **“Bildungsstandards” from the KMK** (‘Kultusministerkonferenz’, Engl. conference of education ministers). This section gives an overview of these two frameworks, what their goal is, what they describe and to what extent they are related to each other.

2.1.4.1. COMMON EUROPEAN FRAMEWORK OF REFERENCE FOR LANGUAGES (CEFR)

The Common European Framework of Reference for Languages (CEFR) was developed by the Council of Europe (COE) and published in 2001 with the aim of “diversifying and intensifying language learning in order to promote plurilingualism in a pan-European context” (CEFR 2001: 4). With evaluation and assessment on the rise in education, the goal of the CEFR was to set objectives and standards on a European level which count for all member state countries.

There are **three main functions** the CEFR fulfills (see CEFR 2001: 1):

“The Common European Framework ...

- ... provides a common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, textbooks, etc. across Europe.
- ... describes in a comprehensive way what language learners have to learn to do in order to use a language for communication and what knowledge and skills they have to develop so as to be able to act effectively. [...]
- ... defines levels of proficiency which allow learners’ progress to be measured at each stage of learning and on a life-long basis.” (CEFR 2001: 1; enumeration added)

Thus, the CEFR is a source for teachers, teacher educators, syllabus designers and education experts. It can be used on a meta level in order to set education guidelines by education representatives of the member countries and also in an applied way by teachers and teacher educators. It can be used by teachers or teacher educators for language assessment and to some extent also as a basis for designing

teaching material. The CEFR defines objectives of language learning and thus describes levels of language proficiency.

The overall **approach of the CEFR** can be summarized as follows:

“Language use, embracing language learning, comprises the actions performed by persons who as individuals and as social agents develop a range of **competences**, both **general** and in particular **communicative language competences**. They draw on the competences at their disposal in various contexts under various conditions and under various **constraints** to engage in **language activities** involving **language processes** to produce and/or receive **texts** in relation to **themes** in specific **domains**, activating those **strategies** which seem most appropriate for carrying out the tasks to be accomplished. The monitoring of these actions by the participants leads to the reinforcement or modification of their competences.” (CEFR 2001: 9; highlighting in original)

The CEFR promotes an **“action-oriented approach”** in the sense that competences describe what learners are able to do or should be able to do. The learner is seen as a **“social agent”** (CEFR 2001: 9) and should act effectively in situations and find its role in society. Competences is the key word in learning and developing specific competences is the ultimate goal in language learning. Competences will be discussed in section 2.1.5., where we take a closer look at what competences are and what kind of different competences the different frameworks suggest and what this implies for language teaching. However, what can be stated is that besides the competences mentioned – general and communicative competence – language activities and strategies play an important role when it comes to language proficiency.



FIGURE 7: FACTORS OF LANGAUGE PROFICIENCY

The goal of language learning and language proficiency according to the CEFR is described as follows:

The learner of a second or foreign language and culture does not cease to be competent in his or her mother tongue and the associated culture. Nor is the new competence kept entirely separate from the old. The learner does not simply acquire two distinct, unrelated ways of acting and communicating. The language learner becomes **plurilingual** and develops **interculturality**. The linguistic and cultural competences in respect of each language are modified by knowledge of the other and contribute to intercultural awareness, skills and know-how. They enable the individual to develop an enriched, more complex personality and an enhanced capacity for further language learning and greater openness to new cultural experiences (CEFR 2001: 43).

It is clear that communicative and intercultural competence need to be established so that the learners become plurilingual and pluricultural, i.e. they draw on their already existing competences and build upon them. Communicative competence (see section 2.1.5.1.) and intercultural competence (see section 2.1.5.2.) also influences the learner's general competences and their personality.

Assessing language proficiency

Communicative language competence can only be established through exercising tasks and performing language activities, which in turn means that learners are in need of these tasks, language activities and strategies in order to build upon their competences. This process needs to be monitored

and evaluated with the help of this framework. As far as **language assessment** is concerned, the CEFR defines three main levels of proficiency with two sub-categories each as depicted in figure 8 (CEFR 2001: 23).

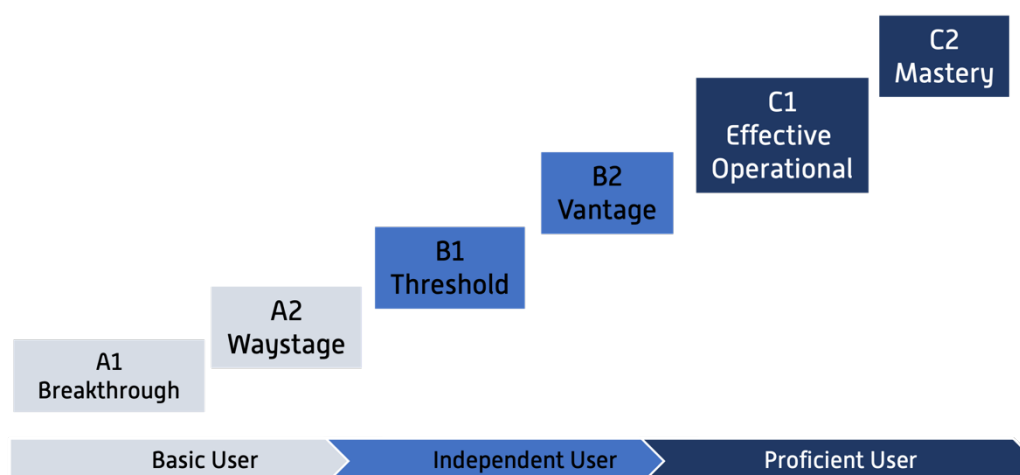


FIGURE 8: CEFR REFERENCE LEVELS

Level A1 (Breakthrough) and A2 (Waystage) describe a basic user, B1 (Threshold) and B2 (Vantage) describe an independent user and C1 (Effective Operational) and C2 (Mastery) are the highest levels of language proficiency and stand for a proficient user (CEFR 2001: 23). Language learners usually start at the bottom with A-level proficiency and make their way up to B-level or C-level proficiency as independent or proficient user. These levels fulfill two main purposes: language assessment and describing objectives. The levels are commonly known from language proficiency tests like TOEFL or Cambridge certificate, where they are used to assess language competence for studying abroad or for specific job needs. Students or applicants usually take a test and get a certificate which testifies their language proficiency with one of the above levels of proficiency.

Basis of these assessment tests and the proficiency levels are competences and the objectives defined in the CEFR. In order to reach a level, the CEFR describes what the students should be able to do. This “can-do-approach” (CEFR 2001: 25) puts the learner into the center of the learning process and so-

called “can-do-descriptors” are defined for each level of proficiency and for each language activity (reception, production, interaction and mediation).

The **descriptors** for A2, B2 and C2 of spoken interaction are displayed in the following table in order to get a better understanding of what the difference between the levels is and thus what the learning progress includes (CEFR 2001: 74).

TABLE 2: DESCRIPTORS FOR SPOKEN INTERACTION

A2	B2	C2
I can communicate in simple routine tasks requiring a simple and direct exchange of information on familiar topics and activities. I can handle very short social exchanges, even though I can't usually understand enough to keep the conversation going myself.	I can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible. I can take an active part in discussion in familiar contexts, accounting for and sustaining my views.	I can take part effortlessly in any conversation or discussion and have good familiarity with idiomatic expressions and colloquialisms. I can express myself fluently and convey finer shades of meaning precisely. If I do have a problem I can backtrack and restructure around the difficulty so smoothly that other people are hardly aware of it.

Comparing these levels makes it very clear that learning is a process and that knowledge and skills build up on each other when progressing from one level to another. While simple and familiar tasks and topics and short exchanges as well as a rough understanding of a conversation are characterized as level A2, fluency, as well as the degree of understanding and action in conversations increases over level B2 to level C2. Teachers as well as learners can easily relate to the different levels of proficiency with the help of the descriptors and they can check their progress as they learn.

In Germany, at the end of secondary education and with the A-level exam ('Abitur'), students are usually at B2 or C1 level (KMK 2012: 14), which is described in the 'global scale' as follows:

[The learner] Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. [The learner] Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. [The learner] Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options (CEFR 2001: 24).

One of the first version of the CEFR dates from 2001 and since then, there have been two major updates. A preliminary version of the updated CEFR was published in 2018 as “Common European Framework of Reference for Languages: Learning, Teaching and assessment: Companion Volume with New descriptors”. The 2018 CEFR Companion volume adds and elaborates on **signing competences** and **plurilingual and pluricultural competences**. Plurilingual and pluricultural competence has become even more important with increasing globalization. These are for example described as knowing and being able to switch between different languages and dialects and being able to make sense of words from other languages or paralinguistics such as gestures and facial expressions (CEFR 2001: 13).

The new CEFR Companion volume itself was published in 2020 and it “updates the CEFR 2001, the conceptual framework of which remains valid” (CEFR 2021: 3). As the following screenshot from the 2020 version of the CEFR shows, new descriptors were added (see blue circles which were added here):

	In the 2001 descriptive scheme	In the 2001 descriptor scales	Descriptor scales updated in this publication	Descriptor scales added in this publication
Communicative language activities				
Reception				
Oral comprehension	✓	✓	✓	
Reading comprehension	✓	✓	✓	
Production				
Oral production	✓	✓	✓	
Written production	✓	✓	✓	
Interaction				
Oral interaction	✓	✓	✓	
Written interaction	✓	✓	✓	
Online interaction				✓
Mediation				
Mediating a text	✓			✓
Mediating concepts	✓			✓
Mediating communication	✓			✓
Communicative language strategies				
Reception	✓	✓	✓	
Production	✓	✓	✓	
Interaction	✓	✓	✓	
Mediation				✓
Plurilingual and pluricultural competence				
Building on pluricultural repertoire	✓			✓
Plurilingual comprehension	✓			✓
Building on plurilingual repertoire	✓			✓
Communicative languages				

FIGURE 9: UPDATES OF DESCRIPTORS (SCREENSHOT FROM CEFR 2020: 23)

It is interesting to see that the descriptors for most of the language activities were established already in 2001, including reception, production and interaction. As far as interaction is concerned, descriptors for online interaction and online interaction itself was only mentioned in the 2020 version. Additionally, descriptors for mediation as language activity have only been added in the new version as well as descriptors for mediation strategies. In general, descriptors for mediation – as communicative language activity or strategy – have only been established in the 2020 version of the CEFR. Another important addition of descriptors was made in the section for plurilingual and pluricultural competence. This means that plurilingual and pluricultural competence had been defined as a goal or competence already in 2001, but only more than 20 years later, descriptors have been added and thus make it more clear what the learners should be able to do.

In sum, it can be stated that the latest version of the CEFR follows the need for a stronger and more sophisticated intercultural communicative competence as it includes new descriptors for mediation and plurilingual/pluricultural competence because “[they] have become increasingly relevant over the past 20 years” (CEFR 2020: 23). A totally new addition are descriptors for “[...] other communicative activities such as online interactions, using telecommunications [...]” (CEFR 2020: 23), which were not mentioned in the 2001 version at all. As these descriptors were just recently added, there is not a whole lot of teaching material available which reflects on these trends and this work tries to close this gap with the help of authentic language data.

This change of focus from communicative aspects to more intercultural aspects, are mainly due to globalization and new ways of collaborating in every day or business life. It is good that the new version of the CEFR has been further developed with an emphasis on communicative and intercultural competences and that it adapts to actual changes and real-world challenges. However, much work needs to be done in order to keep up with these challenges and to adapt language teaching to this “new world” almost 20 years later. To make it a little bit more clear how fast this transformation has been: In 2001, Windows XP had just been launched by Microsoft and many people were using it for many years, but it is very much outdated today. The latest update to the CEFR needed almost twenty years and within that time there have been many Microsoft updates as our behavior, our computer use and our whole (technological) life has changed over the years, for example in terms of (online) collaboration. Of course, the main objectives and frameworks need to be stable over a longer period of time, but language teaching has to keep up with this (digital) transformation and also needs more “updates” in between. In language teaching, we should for example also focus more on collaborative aspects; it has to shed light on how people communicate with each other online. And the latest aspects including Artificial Intelligence (AI) are not even mentioned in this version of the CEFR. This shows, that there is constant change and language learning needs to adapt to it and set the pace in going forward.

The Council of Europe emphasize that the CEFR is a piece of advice and “does not offer ready-made solutions but must always be adapted to the requirements of particular contexts” (CEFR Website³). This is quite obvious for political boundaries and standards, but this is also the biggest issue. It is up to each member state country to what extent they implement these standards into their own education policy and it is also up to the teachers how to put these objectives into content and methods or simply spoken how these guidelines find their way into teaching and learning in the classroom. It is freedom of teaching on the one hand but this also means that the application is dependent on the teachers’ motivation, creativity and responsibility and there comes much variation into play. Examples, support and accompanying material which is ready-to-use for the teachers is not produced to a large extent. The **FREPA** (see section 2.1.5.2.4.) offers some hands-on material, but it is still not sufficient and in my regard not very much known and used in the classrooms. Many publishers and school material editors are still focusing on language activities and communicative competence. Material for improving intercultural competence is very hard to find. This is also one main rationale for this dissertation project: Providing examples and authentic material which can be used in the classroom when it comes to improving communicative and intercultural competence.

2.1.4.2. EDUCATION STANDARDS IN GERMANY

The competences mentioned above from the CEFR and the guidelines as well as the objectives were adopted by the different European union member countries. In Germany, the CEFR was the basis for the **education standards** for second languages. In Germany, the “Kultusministerkonferenz” (abbreviation: KMK, which stands for conference of ministers of education and culture) is responsible for setting education standards – for every school subject and school level – which form the basis for

³<https://www.coe.int/en/web/common-european-framework-reference-languages/uses-and-objectives>
(Accessed 19 November 2020)

the different curricula in each federal state. The teachers then take the state curricula as the basis for their teaching in the EFL classroom (see figure 10).

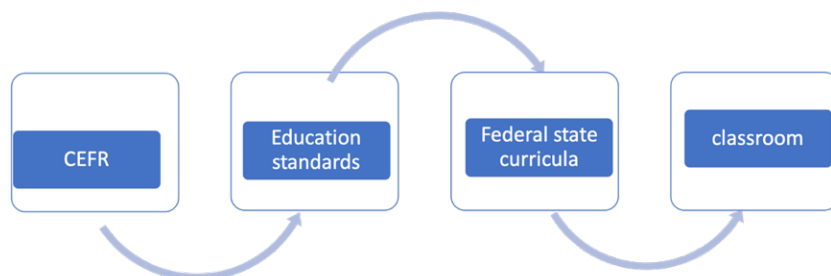


FIGURE 10: FROM CEFR TO CLASSROOM

The education standards have three main objectives: (1) guaranteeing transparency of expectations, (2) contributing to competence-based teaching and (3) forming a basis for assessment (see KMK 2012: 5). The education standards for English at secondary level in German schools date from 2012, when the KMK passed these decisions⁴. They are established especially for secondary schools with the aim of reaching the “Allgemeine Hochschulreife” (high-school diploma, Abitur) and orient to **B2 level** of the CEFR (KMK 2012: 14), the so-called **independent user (vantage)**:

Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options (CEFR 2001: 24).

According to the education standards, this level of proficiency should be achieved through discussing topics, which are relevant to the students (KMK 2012: 12):

- Topics of life and experience of the world of adolescents

⁴https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2012/2012_10_18-Bildungsstandards-Fortgef-FS-Abi.pdf (Accessed 19 November 2020)

- Topics of public life and cultures
- Topics of every-day and professional life
- Topics of global importance

Besides these fundamental decisions what should be content of the lessons, the KMK lists **five competences** in their competence scheme for teaching English in secondary schools:

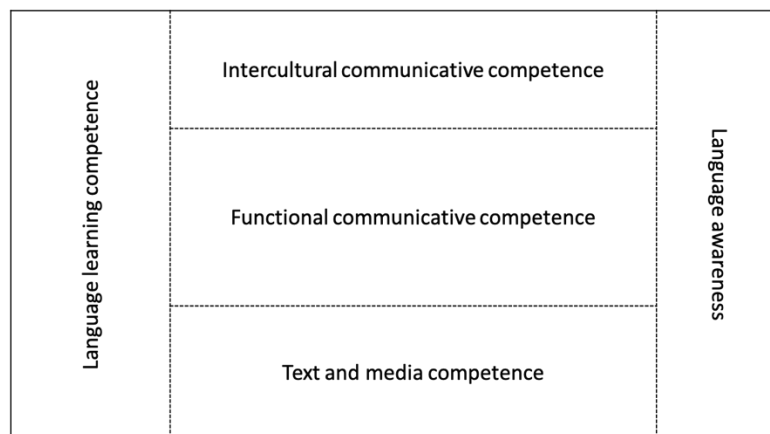


FIGURE 11: COMPETENCES ACCORDING TO THE GERMAN EDUCATION STANDARDS (KMK 2012: 12)

These five competences in the German education standards – intercultural communicative competence, functional communicative competence, text and media competence, language learning competence and language awareness – partially overlap with the competences in the CEFR. Language learning competence and language awareness – partially overlap with the competences in the CEFR. Language learning competence, language awareness and text and media competence can be subsumed under ‘general competences’. It is also important to note that language learning competence and language awareness draws on the competences in the middle, as they are basic in language learning and affect intercultural as well as functional communicative and text and media competence. These basic competences, however, will not be further discussed here, as the focus lies on communicative and intercultural competence. It is important to mention that all lines in figure 11 are dotted – as in the original figure from the KMK – which should imply that the competences do overlap and are intertwined.

With the introduction of these two frameworks and the (brief) description of competences, let us now have a closer look at what competences are, how they are defined in the respective frameworks and what implications for teaching arise.

2.1.5. COMPETENCES

With evaluation and assessments rising in education and thus with the shift in didactics from **focusing on input to focusing on output** (see section 2.1.3.3.), standards and benchmarks have become more and more important. Focusing on input stressed the content what should be taught and curricula listed several elements which should be included into teaching. Focusing on output highlights **competences**. Competences are therefore at the heart of language teaching.

There is not one clear **definition of competences**. The word ‘competere’ is Latin and can be translated as meeting, combining or being able to do something. One interpretation is that requirements of a certain situation meet with the person’s skills and the person has the competence to handle this situation. In linguistics, there has been a different take in the 1950s, as there was a dichotomy between competence and performance, which was mainly coined by Noam Chomsky (Chomsky 1965; cited in Mukherjee 2002: 78). Competence in this context meant having knowledge about a language and the language system, whereas performance has an applied part and meant that language needs to be used in context and in speaking with others in order to fully “know” a language. There has been growing critique on this dichotomy especially in applied linguistics.

The definition of competence from an educational or methodological point of view is different: Performance is included in the competence definition as well as personal attitude. Erpenbeck (2012: 14) came up with a model of four key competences that is often used in educational science: personal, professional-methodological, social-communicative as well as activity and implemented-oriented

competence. In anglophone educational research, the **KSAO model** is often used. In the KSAO-model, competence consists of four factors: **knowledge, skills, abilities and other characteristics**.

Weinert's definition is usually used in the field of language education in Germany:

Competences are the cognitive abilities and skills available to or learnable by individuals to solve specific problems, as well as the associated motivational, volitional and social willingness and skills to use the problem solutions successfully and responsibly in variable situations (Weinert 2001: 27, my translation).

Thus, competence, as it is also understood in this work, is more than knowledge. It is a combination of knowledge, skills, behavior and the ability to (re)act in certain situations. Similarly to Weinert's definition, competences are defined in the CEFR as "the sum of knowledge, skills and characteristics that allow a person to perform actions" (CEFR 2001: 9). It is important to note that competences are not only declarative knowledge, but also procedural knowledge, "know-how", and social skills and characteristic traits. And this means that the way of teaching – with focusing on input – consequently changed towards focusing on output and thus on competence. Acquiring declarative knowledge is not the highest goal any more. Skills and know-how have become an essential part in learning and thus teaching needs to adapt to this in a way that skills and know-how need further practice and applications.

Competence-based teaching thus is a combination of acquiring knowledge and providing situations where learners can adapt and apply their knowledge. In that case, the students see to what extent they can apply knowledge and act in specific situations, which is the general aim of teaching. If the application of knowledge is not successful, the most important part of the learning process is not achieved. This means that the learners might either want to revise the facts and declarative knowledge or practice more in order to work on their procedural knowledge. Application needs practice and this varies across learners due to their individual learning potential. Only through continuous practicing

and learning from mistakes⁵, the learners' full competence can be acquired and the full learning outcome is visible.

In this context, it is important to mention – briefly – the term **differentiation**. Learners have different competences and different abilities to learn. This leads to different speed of learning, i.e. some learners will need more practice while other learners are faster and do not need to revise or practice too much. Of course, this also depends on the learners' mindset, the content and their motivation, but leaving these personal and volatile competences aside, it is important as a teacher to be aware of this fact. Therefore, teachers need to adjust to these different competences and to provide material for the learners so that everyone has the same chance to achieve full competence whether they need more time or are fast.

Two main competences of language learners, which are listed in the CEFR in the 2001 version, are **general competence** and **communicative language competence**. According to the CEFR, “[g]eneral competences are those not specific to language, but which are called upon for actions of all kinds, including language activities” (CEFR 2001: 9).

General competences of a language learner include (CEFR 2001: 12):

- **knowledge** (*savoir*), as in declarative knowledge like empirical and academic knowledge;
- **skills and know-how** (*savoir faire*) as in procedural knowledge;
- **existential competence** (*savoir-être*), which entails characteristic traits and attitudes;
- and **ability to learn** (*savoir apprendre*) also in the sense of “knowing how, or being disposed, to discover “otherness” – whether the other is another language, another culture, other people or new areas of knowledge”.

These general competences are not subject or domain specific, but rather have to do with the learners themselves. They include personal and social competences as well as methodological competences, which could be used in all subjects at school or different domains of the learners' lives.

⁵ Note: Mistakes in this context are very positive. Mistakes are not seen as deficiencies but rather as chances or situations to learn from.

The German education standards list five competences (see figure 11): Language learning competence and language awareness as basic competences and text and media competence, intercultural communicative competence and functional communicative competence. It is interesting to see that the KMK further divided communicative competence into ‘intercultural communicative competence’ and ‘functional communicative competence’ and thus gave special value to each of these two competences. Language learning competence, language awareness and text and media competence fall into the “general competences” category of the CEFR.

2.1.5.1. COMMUNICATIVE COMPETENCE

Before looking at how communicative competence is defined in the CEFR and in the education standards, we must take a step back and look at how the term ‘**communicative competence**’ itself evolved. In the 1950s in the field of linguistics, there existed a dichotomy between competence and performance (see Chomsky 1965; cited in Mukherjee 2002: 78). Competence was defined as the abstract knowledge about languages, whereas performance meant the actual language use (see Mukherjee 2002: 78). Applied linguists have criticized this view and argued that competence also needs performance, i.e. not only having knowledge about a language is enough, but also being able to use a language is necessary (Mukherjee 2002: 78). Hymes (1972), a socio-linguist, coined the term ‘communicative competence’, and thereby extended Chomsky’s definition of linguistic competence, with a social component. Hymes suggests four questions for language use:

1. Whether (and to what degree) something is formally *possible*;
2. Whether (and to what degree) something is *feasible* in virtue of the means of implementation available;
3. Whether (and to what degree) something is *appropriate* (adequate, happy, successful) in relation to a context in which it is used and evaluated;
4. Whether (and to what degree) something is in fact done, actually *performed*, and what its doing entails. (Hymes 1972: 281)

With these questions, Hymes combines the concepts of competence (possible and feasible) and performance (appropriate and performed) overcoming the dichotomy, i.e. he adds a language in use aspect to the formally linguistic and theoretical definition. “Speakers need to know not only grammatical structures but also norms of usage and appropriacy in a given social context” (Hymes 1972, cited in Savignon 2018: 2). Thus, he stressed appropriateness and the socio-cultural context of an utterance and highlights actual language use.

Canale and Swain (1980) introduced four elements of communicative competence: linguistic competence, discourse competence, strategic competence and sociolinguistic competence. They included elements of sociolinguistics, discourse studies and pragmatics into their definition and started to look at the context and co-text of an utterance and not only at individual words. Van Ek (1986) even developed communicative competence further and added two more elements to Canale and Swain’s definition: socio-cultural competence and social competence. These two give a cultural note to communicative competence. According to van Ek, speakers need to function in several cultures (socio-cultural competence) and need to be familiar with customs in social groups and need to communicate with other people and show empathy (social competence) (van Ek 1986). Communicative competence developed from a more linguistic or lexical definition over socio-linguistic and pragmatic elements to social and intercultural aspects. This already shows, that context is always very important in communication and that culture and communication are very closely related and intertwined.

In this work, I try to separate communicative competence and intercultural competence in order to gain a deeper understanding of each individually. In the analysis, we will look at communicative competence and aspects of intercultural competence in the known standards and frameworks individually and conduct the analysis and application for both competences. The idea is that elements of both competences combined result in a combined competence – intercultural communicative competence – with the intercultural speaker as the goal in language teaching. The following sections provide an overview of how communicative competence is defined in the CEFR and the German education standards.

2.1.5.1.1. COMMUNICATIVE COMPETENCE IN THE CEFR

According to the CEFR, “language learning comprises the actions performed by persons who as individuals and as social agents develop a range of competences, both general and in particular communicative language competences” (CEFR 2001: 9). This means that besides the so-called general competences, the second important competence in language learning is communicative competence. Communicative competence in the CEFR is described as follows: “Communicative language competences are those which empower a person to act using specifically linguistic means” (CEFR 2001: 9) and consist of (1) **Linguistic competence**, (2) **sociolinguistic competence** and (3) **pragmatic competence** (CEFR 2001: 13).

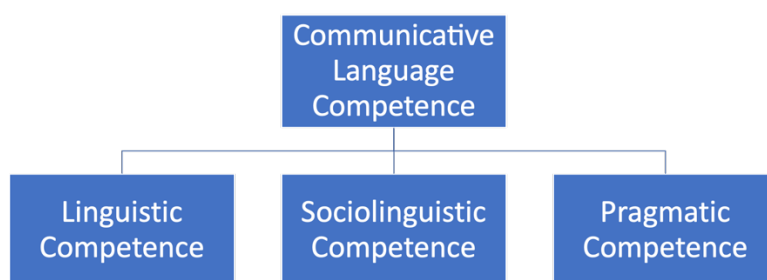


FIGURE 12: ASPECTS OF COMMUNICATIVE LANGUAGE COMPETENCE IN THE CEFR

Linguistic competence is based on lexical, syntactical and phonological skills, while sociolinguistic competence is described as language in use and being able to follow socio-cultural norms. Pragmatic competence refers to speech acts and discourse in general (CEFR 2001: 13).

These three competences are again divided up and further explained as follows:

Linguistic competence	Socio-Linguistic competence	Pragmatic competence
• Lexical competence (fixed expressions, single word forms) • Grammatical competence (morphology, syntax) • Semantic competence (lexical semantics, grammatical semantics, pragmatic semantics) • Phonological competence • Orthographic competence • Orthoepic competence	• Linguistic markers of social relations (greetings, address forms etc.) • Politeness conventions • Expressions of folk wisdom • Register differences • Dialect and accent	• ‘Discourse competence’ (organised, structured and arranged) • ‘Functional competence’ (used to perform communicative functions) • ‘Design competence’ (sequenced according to interactional and transactional schemata)

FIGURE 13: COMMUNICATIVE LANGUAGE COMPETENCES IN THE CEFR (CEFR 2001: 108F)

In acquiring these competences (see figure 13), communicative language competence is established. Language proficiency and “progress in language learning is most clearly evidenced in the learner’s ability to engage in observable language activities and to operate communication strategies” (CEFR 2001: 57). Thus, language activities and communication strategies also play an important role when it comes to language proficiency. Establishing communicative language competence and enriching language proficiency can only be reached through **communicative activities and strategies**.

Language activities have to do with learners “processing (receptively and/or productively) one or more texts in order to carry out a task” (CEFR 2001: 9). Examples of language activities are **reception** and **production** activities, as well as **interaction** and **mediation**.

Reception activities include (silent) reading of any kind of written works such as the news or scientific studies and listening for example when following the media (CEFR 2001: 14). This means oral, audio-visual and reading comprehension in general (CEFR 2020: 47). Oral comprehension for example is understanding conversations (between other people), understanding as a member of audience, understanding announcements and instructions as well as understanding media or recordings (CEFR 2020: 47).

Production activities can be distinguished between oral and written production, which is basically speaking and writing. Oral production activities include monologues when for example describing experience, giving information or putting a case, public announcements or addressing audiences when for example giving presentations. Written production includes writing in general, creative writing and writing reports or essays (CEFR 2020: 61). Usually, these kind of activities are evaluated, which means that there are other people who judge or even grade the products, meaning a presentation or a report (CEFR 2001: 14).

Interaction happens when a second person is involved and people talk or write to each other. This exchange of information or facts – back and forth – either in a written exchange or a conversation are central to interaction. Usually, reception and production activities alternate or even happen simultaneously, i.e. people speak and listen to each other. Interlocutors take turns according to turn-taking rules and basic pragmatic principles and sometimes utterances can also overlap (CEFR 2001: 14). Interaction can be distinguished between oral interaction, written interaction and – new in the 2020 version of the CEFR – online interaction. Oral interaction includes understanding an interlocutor, conversation, informal and formal discussions, goal-oriented co-operation, obtaining goods and services, information exchange, interviewing and being interviewed as well as using telecommunications (CEFR 2020: 71). As far as written interaction is concerned, the CEFR lists overall written interaction, correspondence, notes, messages and forms as formats (CEFR 2020: 71). Interaction has traditionally focused on face-to-face interaction when people actually meet and talk to each other. With new online collaboration tools on the rise, **online interaction** has also become an important part of interaction. Online interaction can be any kind of exchange online, for example writing an E-Mail, reading a blog post or taking part in an online meeting and thus draws on all categories above. Online communication has only been part of the 2018/2020 version of the CEFR and only in that version, the descriptors have been established (see figure 9; see also section 2.1.4.1.).

Mediation has for a long time been understood as translating or interpreting, which is only one aspect. Key to mediation in language learning context is that people can understand each other language-wise

as well as culture-wise. When people speak different languages, the mediation task could be translating or interpreting as well as formulating a summary of what has been said. If people stem from different cultural backgrounds, it is also about explaining cultural concepts and creating common ground so that the exchange of these people can still take place. Mediation in a non-linguistic way has to do with overcoming a conflict between two people when a third person – a mediator – tries to work towards mutual understanding. The CEFR in the 2001 version focused on the communicative and intercultural mediation, even though the descriptors for mediation in online interaction and for intercultural communication have only been established in recent years for the publication of the 2020 (and 2018) version. The definition has been further developed and mediation has gained further importance: Mediation – be it communicative, interpersonal or intercultural – is not only seen as language activity any more but also as an approach and objective (COE 2016: 58).

Mediation activities in the 2020 version of the CEFR include mediating a text, concepts and communication (CEFR 2020: 90). Mediating a text includes relaying specific information, explaining data, processing text, translating a written text, note-taking, expressing a personal response to creative texts and the analysis of criticism of creative texts. Mediating concepts could take place when somebody is collaborating in a group, like when facilitating collaborative interaction with peers or collaborating to construct meaning (CEFR 2020: 90). Or it could happen when leading group work, for example when managing interaction or encouraging conceptual talk (CEFR 2020: 90). Mediating communication is described as facilitating pluricultural space, acting as an intermediary and facilitating communication in delicate situations and disagreements (CEFR 2020: 90).

Usually, these language activities are carried out in a strategic way, meaning that language or **communicative strategies** underlay these activities. Communicative strategies are therefore seen as a binding element between competences and activities. **Language strategies** are “any organised, purposeful and regulated line of action chosen by an individual to carry out a task [...]” (2001: 9). This task can either be self-initiated or other-initiated, which means that the students could either chose to fulfill a task or they could be confronted with a special task by the teacher. Communicative strategies

are meta-cognitive processes which happen before, while and after the language activity. These strategies follow a three-step which is usually carried out as a line of tasks: plan (1), execute (2), monitor and repair (3). The descriptors were not established for each strategy in the 2001 version of the CEFR, the ones in table 3, which are written in italics were only added in the 2020 version.

	Reception	Production	Interaction	Mediation
Planning	<i>Framing</i>	Planning	N/A	*
Execution	Inferring	Compensating	Turn-taking Co-operating	Linking to previous knowledge Adapting language Breaking down complicated information Amplifying a dense text Streamlining a text
Evaluation and Repair	<i>Monitoring</i>	Monitoring and self-correction	Asking for clarification <i>Communication repair</i>	*

TABLE 3: COMMUNICATIVE LANGUAGE STRATEGIES IN THE CEFR (CEFR 2020: 35)

As far as **production strategies** are concerned, people usually plan what they are going to say or write. While they are producing an utterance (execution), they usually compensate about what the interlocutor will say as well as about what they will answer. After the utterance has been made or the text has been produced, people should monitor what was written or said and might end up in self-repair if they want to correct themselves. It is important to note here, that the CEFR decided to mention only execution strategies for mediation and leave out planning and evaluation and repair, even though these steps could also be possible in mediation activities. The CEFR (2020) lists for **reception strategies** only identifying cues and inferring (CEFR 2020: 47). As for **interaction strategies**, the CEFR mentions “turntaking, co-operating and asking for clarification” (CEFR 2020: 71). **Mediation strategies** can be distinguished in “strategies to explain a new concept” or “strategies to simplify a text” (CEFR 2020: 90).. In order to explain a new concept, learners could link to previous knowledge,

adapt language or break down complicated information. In order to simplify a text, learners could amplify a dense text or streamline a text (CEFR 2020: 90).

The goal of using these language strategies is to be effective in communication or as the COE puts it: “[to] fulfil the demands of communication in context and successfully complete the task in question in the most comprehensive or most economical way feasible depending on his or her precise purpose” (CEFR 2001: 57).

2.1.5.1.2. FUNCTIONAL COMMUNICATIVE COMPETENCE IN THE GERMAN EDUCATION STANDARDS

The German education standards differentiate between **functional communicative competence** and **intercultural communicative competence**. Functional communicative competence is further divided into five central “kommunikative Fertigkeiten” (KMK 2012: 12), hence referred to as **communication activities**:

- Listening and audio-visual comprehension
- Reading comprehension
- Speaking / Interaction (expressing oneself as well as taking part in conversations)
- Writing
- Mediation

As far as **listening and reading comprehension** is concerned, learners should be able to understand the main message of a text – written or spoken – and excerpt and thematically categorize specific information. They should also be able to understand implicit messages (KMK 2012: 12). Speaking is divided into two categories, namely taking part in conversations and uttering monologues. Learners who **take part in conversations** should be able to “express themselves fluently, correctly and appropriately to people and situations. They should be able to interact in conversations even in situations when the topics are fairly unknown to them” (KMK 2012: 16). **Monological speaking**

includes “giving presentations, expressing a point of view as well as explaining different options, for example advantages and disadvantages” (KMK 2012: 16). **Writing** is described as “composing texts of personal or professional interest appropriately to people and specific to different text types. The learners should be aware of techniques and strategies for formal, informal and creative writing” (KMK 2012: 17). **Mediation** means that the learners should be able “to reproduce authentic, written or spoken texts in the respective other language” (KMK 2012: 18).

These communicative skills relate to the communicative language activities and strategies in the CEFR (reception, production, interaction and mediation). Listening and audio-visual comprehension as well as reading comprehension could be subsumed under reception activities in the CEFR. Speaking and writing would belong to production activities and mediation exists as an own category in the CEFR and the education standards. It is interesting that the CEFR lists interaction as a separate activity, but the education standards only mention interaction together with speaking, i.e. a production activity. It is clear that the boundaries are very fuzzy here and that productive activities usually happen together with interaction. However, to be clear, it would have been better to put interaction as a separate category as there also arise different strategies for interaction and thus speaking and interaction has to be taught individually. Interaction is harder to teach than production activities like (monological) speaking and writing. In interaction, receptive and productive activities come together, learners have to pay attention to their interlocutor and react spontaneously in the situation, whereas in productive activities, the planning phase is usually more elaborated. Collaboration and communication belong to the “future skills” and are extremely important in a work environment and in everyday communication. Therefore, interaction activities and interaction strategies become more and more important and need to be taught and learned. The EFL classroom is the best location to do that as communication (in other languages) and collaboration can be practiced together.

For each of these activities, the KMK suggest two levels – a basic level (“Grundlegendes Niveau”) and an advanced level (“Erhöhtes Niveau”) with descriptions what the students should be able to do (Can-

do approach). These descriptions are – compared to the CEFR – not very elaborated and leave room for interpretation. They are rather objectives than descriptors.

One of the objectives in the category “speaking / interaction” for example is: “The students can use adequate communicative strategies consciously in order to deal with non- or misunderstanding” (KMK 2012: 16, my translation). The problem here is that it is not further described what these communicative strategies are or how the students should deal with misunderstandings. It is expected that they have certain strategies to deal with it. However, it is not easy for teachers to know what kind of strategies should be taught and how they could be practiced (and tested) in class.

In addition to these skills mentioned above, the KMK lists another two central elements of functional communicative competence: **linguistic resources** (“sprachliche Mittel”) and **communicative strategies** (KMK 2012: 12). While linguistic resources already belong to communicative competence in the education standards for “Mittlerer Bildungsabschluss” (intermediate diploma, usually obtained after grade 9), communicative strategies are added in the education standards for “Allgemeine Hochschulreife” (general high school degree, after grade 12 or 13).

Linguistic resources (“sprachliche Mittel”) include the following (KMK 2012: 15):

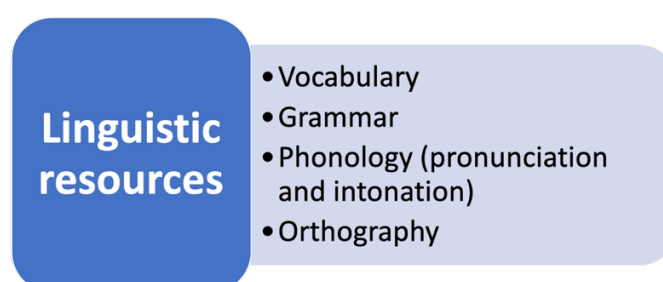


FIGURE 14: LINGUISTIC RESOURCES

It is important to note that these linguistic resources have “serving function as the ultimate goal is successful communication” (KMK 2012: 15). In order to take part in a conversation, learners have to know words (vocabulary). They have to know how these words are used in a sentence (grammar), how

they are pronounced (phonology) or spelled (orthography). Thus, linguistic resources build a repertoire which the learner should make use of while speaking or writing for example. These linguistic resources form the basis and should help the learner to utter or understand a sentence. In order to follow or take part in conversations, the learners do not only need to know specific words or how these words are pronounced, they also need communicative strategies. However, communicative strategies are not further developed in the education standards. According to the KMK, they are mentioned in the description for the respective communicative skills (see KMK 2012: 15), but I would argue that these strategies need to be addressed as a separate aspect in order to pay tribute to their importance. Communicative strategies are also further established in the CEFR and are – besides communicative language activities – a crucial factor in developing communicative competence.

2.1.5.2. INTERCULTURAL COMPETENCE

“Culture is in language, and language is loaded with culture” (Agar 1994: 28). Language and culture are closely related and intertwined and should both be subject to teaching (see Kramsch 1991, 1998a). Intercultural competence is one of the main competences in EFL teaching according to the CEFR and the German education standards. This section looks at how intercultural competence is defined in literature and by the European and German governments. Before looking at intercultural competence, some important aspects regarding culture in general and some main concepts will be described first.

2.1.5.2.1. WHAT IS CULTURE?

Before taking a look at intercultural competence, it is important to briefly focus on “**culture**” itself. There is not one clear cut definition of what culture is and it would go far beyond this work to discuss

the term “culture” in its full depth. However, the basic principles and concepts of culture will be presented very briefly as a basis before looking at intercultural competence as an important competence in EFL teaching.

There are three basic concepts of culture, which were summarized by Lüsebrink (2005) based on the discussions by various researchers during the decades. In the **intellectual aesthetic concept** (1), “education” and “art” play an important role as culture is represented usually by writers, art or composers (Lüsebrink 2016: 10). Moral/ethical as well as aesthetic values are the basis of this concept. The **material** concept (2) goes back to the original meaning of “agricultura”, translating into agriculture or farming. Material aspects of a culture are highlighted in this concept. According to the **anthropological concept** (3), culture is described as “the software of the mind” (Hofstede 1993: 13, see also Lüsebrink 2016: 10) meaning how people are programmed in a group and can therefore be distinguished from another group (Hofstede 1993: 19). This includes ways of thinking, feeling and acting and thus describes collective patterns of a specific society.

In order to classify different cultures, Hofstede (1993: 22f) defined four **levels of cultural differences**:

TABLE 4: CULTURAL DIMENSIONS

Symbols	Symbols are words, gestures, images and objects that have a certain meaning which can only be fully recognized by the people from the same cultures
Heroes / role models	Heroes / role models are people (dead or alive, real or fictional) who represent character traits that are well respected and liked in a culture. Many people within that culture can identify themselves with these role models and they are culturally specific and in the cultural memory
Rituals	Rituals are collective activities that are seemingly superfluous but are necessary for social reasons
Values	Values are the “core” of a culture; i.e. a general tendency of a group to prefer certain circumstances to others, often feelings regarding basic things

One model, that puts these cultural dimensions into a visual figure is the so-called “**onion model**” (see Lüsebrink 2016: 13). In the onion model, symbols represent the outer layer of the onion and values

are the core of the onion; heroes and rituals in between. In order to understand a culture in its deep values, people have to work their way through the different layers of the onion, from symbols, over heroes, norms and rituals to values (Hofstede 1993: 22f; Lüsebrink 2016: 12). In learning about these cultural dimensions, people also learn more and more about the culture itself and end up in the core of the culture.

Cultural dimensions (cultural standards)

“Cultural dimensions” or “cultural standards” consist of shared norms, values, attitudes etc. on a culture and give members of this culture a sense of orientation regarding their own behavior (Thomas 1991: 7). With the help of these cultural dimensions, people can classify their own behavior as normal or typical and even whether it is acceptable in a certain culture (see Thomas 1994: 151). Two major classifications of cultural dimensions should be mentioned here: Hofstede (1993) defined five cultural dimensions including power distance, individualism vs. collectivism, masculinity vs. femininity, uncertainty avoidance index and long-term vs. short-term orientation.

The second classification stems from E.T. Hall (1959), who defined three cultural dimensions. The first dimension E.T. Hall (1959, 1966, see also Hall/Hall 1990) described is high-context vs. low-context cultures, which describes a **communication style** and to what extent people communicate in a direct or indirect way and to what extent harmony and respect for others are considered important. The second dimension has to do with **time** and is described as monochronic and polychronic cultures. In polychronic cultures, punctuality is less important and several tasks are carried out simultaneously. This means more flexibility, creativity and spontaneity in contrast to monochronic cultures, where punctuality is central and careful planning are key. Monochronic cultures are thus said to be structured, reliable, effective and plannable. In the last dimension, cultures can be classified according to **space** and their degree of **social distance**, i.e. whether people prefer a low or high level of social distance. Low distance cultures are characterized through social closeness, whereas high distance

cultures are more objective and focused on work and only sometimes enter private sphere conversations.

It is important to note, that these dimensions mentioned above are **continua**, describing tendencies of specific groups or cultures and which count usually for the majority of people within one culture, but not for every individual. These dimensions can help in intercultural situations or in intercultural learning in order to have a better understanding of the interlocutor. It is very important to use cultural dimensions in a very sensitive way in order to prevent stereotypes and misunderstandings. Taking them as an orientation or as a starting point in the intercultural understanding process can be helpful. However, not only culture, but also the person itself and the situation – often referred to as the PSC model – should be taken into account in intercultural encounters. Whenever a person enters an intercultural situation, it is always the person with its characteristics and traits, its cultural background and the situation which should be regarded as a whole.

People from the same social environment share – at least partially – specific cultural phenomena. For a long time, there has been a distinction between “Culture” vs. “culture”. Culture with a capital C or “high” culture draws on the intellectual aesthetic concept and is concerned with literature and art which define a culture. The term culture with a lower case “c” is closer to the anthropological concept of culture and is usually associated with the concept of “doing culture”. Following the renowned Birmingham School of Cultural studies and its fellows like Stuart Hall, E.P. Thompson and Raymond Williams, culture is regarded in a holistic way and defined as a **“whole way of life”** (Williams 1981: 11). In this view, popular culture is also one aspect of culture and needs to be taken into account. When doing cultural studies, the influence and the effect of culture on people’s lives is under investigation and this includes all aspects, not only intellectual or material aspects.

The concept of culture in this work follows the Birmingham School of cultural studies approach and Williams’ definition in looking at culture as a whole and including **all aspects of life**. It follows the anthropological concept of culture and Thomas (1993)’s definition of culture as a “special collective

orientation system” (Thomas 1993: 380) of a group, which we automatically learn growing up in a group and thus gives us a feeling of belonging or identity towards a specific group. It also provides the (unwritten) “rules” for living and working together in a country (Thomas 1993: 380). It is also based on the **constructivist-realist paradigm** (see Hua 2016). In the constructivist paradigm, culture is socially constructed and can only be understood through interaction as understanding a culture is very subjective. The aim in the realist paradigm is explanation (see Hua 2016). Culture is seen as underlying structure and the research focus is on explaining actions through finding this underlying structure (Hua 2016).

Interculturality, transculturality and hybridity

Interculturality focuses on what happens “between” two cultures, while these two cultures are regarded as separate and not overlapping. People have a sense of belonging to one culture and see people from other cultures as “the other” (Lüsebrink 2016: 17). This is a contrastive way of looking at cultures and looking at differences between cultures. This concept is sometimes criticized as cultures seem to be bound to nations, however this is not the case any more in a globalized world. What is clear in today’s globalized world is that cultures are usually not homogenous, but rather dynamic and interactive and can overlap. These more global views on culture are described with the concepts of “**transculturality**” (see Welsch 1990) and “**hybridity**” (see Bhabha 2007). Transculturalism is characterized by plural identities and the interconnection of different cultures as cultural “networks” (see Welsch 1992). Hybridity describes a situative culture or a third space – similar to interlanguage – which arises when people from different cultures interact and negotiate meaning (Bhabha 1994: 86). The focus is not on the difference between these people and their cultural background but on the hybrid culture which arises during their interaction.

Identity and representation

One aspect which is also closely related and intertwined with culture is **identity**. Identity in this work follows the anti-essentialist approach viewing identity as discursively constructed (Baker 2015) and socially negotiated so that “groupness” unfolds (see Edwards 2009). Cultural identity is, according to Friedmann (1994: 29) the “attribution of a set of qualities to a given population”, which means that “personal, sexual, national, social, and ethnic identities [are] all combined into one” (Chen 2006b: 12). Language also plays an important role in the (discursive) construction of identity. There is not only one identity people “own”, but they take on “different identities at different times” (Hall 1992: 277) when communicating with other people, which makes the concept of identity very complex and fluid. People even change positions and voices every day, which is usually referred to as plural identities or the plural “Self” and “Other” (see Baumann 2004, Maffesoli 1996). Closely related to identity is the term social representativeness.

Social representations are systems of values, ideas and practices which enable communication to take place among members of a community by providing them with a code for social exchange and a code for naming and classifying unambiguously the various aspects of their world and their individual and group history (Moscovici 1961: xiii).

This means that sociality is created and social representations allow us to position ourselves within a group or towards the “Other” (Dervin 2012: 185; see also Howarth 2002). Stereotypes and Othering are types of social representations (Dervin 2012: 186) and appear to a large extent in intercultural communication. Especially stereotypes often have a negative connotation and are part of “collective meta-attitudinal” discourse (Moore 2003: 14) and set borders between groups, but these borders could as well be crossed. Stereotypes give an orientation as long as they are treated sensitively and consciously and without judgement. As long as they are descriptions of social groups and not regarded as absolute statements, which are taken in unreflectively, they can serve as examples in intercultural learning. In intercultural encounters, identity is found to be influenced by personal and group identities, situational aspects and contributions by other conversation partners (see Swann & Bosson 2008: 460).

Intercultural communication

When people from different cultures meet, these people usually have to understand each other, i.e. intercultural communication is needed. **Intercultural communication** – from a linguistic perspective – unfolds in situations when people of different cultural backgrounds communicate with each other (Schugk 2004). It is important to note here, that intercultural communication plays an important role and according to Schugk's definition, intercultural communication can also happen within one nation (intranational), if the interlocutors have different ethnic backgrounds.

Lüsebrink uses a broader definition of intercultural communication, which also entails interpersonal interaction as well as mediated intercultural communication, for example "forms of communication in movies, TV, radio, on the internet or other media or other forms of daily intercultural communication across different communication technologies and media" (Lüsebrink 2016: 8). This broader definition is therefore not restricted to language or communication, it also entails the negotiation of meaning and concepts. It is important to note here that Lüsebrink and Schugk's take on intercultural communication is also not only concerned with the difference between these two cultures, but how people communicate with each other and thus negotiate meaning. Their concept of intercultural communication underlays a globalized world with transculturalism and hybrid cultures, even though their focus is in on how people with different linguacultural backgrounds communicate with each other.

Intercultural learning

As far as **intercultural learning** is concerned, Lüsebrink (2005) lists three main abilities the learners should acquire. The first one is to take an outside perspective (1), meaning that the learner should take a step back and look at his/her own views and actions from the outside. The second ability is that the learners should be able to tolerate ambiguity (2), which means that even if differences in interests or needs arise, the learner should face them with respect and try to reach an agreement. The third goal

is that the learners should be able to empathize (3), which means that they should be able or willing to familiarize themselves with unknown situations and unknown environments (Lüsebrink 2005: 68).

According to Bennett (1993), there are two main variables in intercultural learning: the **ability to interact and intercultural sensitivity**, which evolves from experience of cultural differences. With the increase of both factors, the learners become interculturally sensitive and can interact in intercultural situations. This development model of intercultural sensitivity consists of six steps: denial, defense, minimization, acceptance, adaption and integration (Bennett 1993: 182). The first three steps are characterized as “ethnocentrism” (Bennett 1993: 182; see also Barmeyer 2000: 2) and the learner at these stages are not really acting in an intercultural way as they are focusing on their own culture and do not accept other cultures. With passing stages four until stage six, the learners accept cultural differences, they adapt to other cultures without leaving their own culture behind and finally integrate themselves into new cultures. These last three steps can be summarized as “ethnorelativeness” (Bennett 1993: 182; see also Barmeyer 2000: 303). Intercultural learning usually takes place when people move to a new country and culture, but it could also appear when people visit foreign countries and make intercultural experiences. Intercultural learning can also be taught – without visiting a foreign culture or country. A very common and useful method is the use of critical incidents.

Critical incidents are situations or examples where typical misunderstandings unfold because members from different cultures behave in a different way than expected. The interlocutors misinterpret culturally specific behavior due to a lack of knowledge and experience (Barmeyer 2000: 109; Heringer 2004: 219). These critical incidents could be ambivalent, conflictual or puzzling (Fiedler et al. 1971: 97). Critical incidents are frequently used in intercultural trainings or workshops in the business context when people are being prepared to visit a foreign country. Critical incidents can be used to make aware of cultural differences and to practice how these misunderstandings could be overcome. They are not very commonly used at schools even though they have great potential in teaching intercultural sensitivity and communication in order to teach the learners how to act in an

intercultural competent way. Intercultural learning forms the basis to acquire intercultural competence in school contexts.

Culture and language teaching

Claire Kramersch, an influential researcher in applied linguistics, who focuses on culture and language teaching, states that successful communication in EFL teaching is culture-bound and she sees language and culture as inseparable (Kramersch 1991: 218). According to her, “language expresses, embodies and symbolizes cultural reality” (Kramersch 1998a: 3). Language is in a way a means to understand and make meaning of different cultures. As a result of this complex interconnectivity between language and culture, Kramersch suggests that “one must be accounted for in teaching / learning the other” (Kramersch 1998a: 3).

There are several ways on how to integrate cultural elements in language teaching: For a long time in EFL teaching, “Landeskunde” (regional studies) was the aspect which brought culture in the EFL classroom, especially the so-called “target cultures” of native English speaking countries like Great Britain or the U.S. With rising interest in intercultural communication in applied linguistics in the 1990s and onward, the term “Landeskunde” was replaced by intercultural (communicative) competence, which the KMK suggests now as competence in EFL teaching in Germany (see KMK 2012: 12). Elements of “Landeskunde” are still part of intercultural competence, like socio-cultural orientation knowledge, but Intercultural competence is a broader concept with not only knowledge, but also skills and attitudes (see section 2.1.5.2.). Interculturality typically has a contrastive aspect, i.e. two cultures which are being compared and differences are stated. Especially in intercultural teaching and communication, it can be useful to focus on differences and also to address stereotypes in order to promote intercultural learning. It is important, however, to look at the big picture in our globalized world. The focus should not only be on differences, but also on similarities and the negotiation of meaning when people with different linguacultural backgrounds communicate with each other.

It is important that the learners are aware of the concepts of interculturality, but also acknowledge that culture in a globalized world is not nationally bound any more, but that it crosses national borders in the sense that transculturalism and hybridity are the reality. The focus, however in teaching, is usually on intercultural competence in order to differentiate between cultures and look at how intercultural communication in Lüsebrink and Schugk's definition works. This accounts especially at a beginner level. In more advanced levels where globalization is also part of the curriculum, aspects of transculturalism and hybridity should be discussed. The ultimate goal, however, is that learners can act in intercultural situations and thus intercultural competence needs to be established.

2.1.5.2.2. ASPECTS OF INTERCULTURAL COMPETENCE

Intercultural competence, according to Deardorff (2004) is "the ability to communicate effectively and appropriately in intercultural situations based on one's intercultural knowledge, skills, and attitudes" (Deardorff 2004: 194). For Beneke, "intercultural competence is to a large extent the ability to cope with one's own cultural background in interaction with others" (Beneke 2000: 108). Lüsebrink (2005) expands the previous definitions and claims that intercultural competence is the competence to act and communicate in an interculturally sensitive way, as well as the competence to understand intercultural signs (Lüsebrink 2005: 9). Lüsebrink's definition of intercultural communication also contains interpersonal interaction in media-mediated situations (Lüsebrink 2016: 8). This broader definition is also used in the German education standards as they state that "the students can communicate in direct and in media-mediated intercultural situations" (KMK 2012: 19). A very common and very widely-spread definition, which summarizes all important aspects from previous definitions is:

Intercultural competence is the ability to communicate effectively in cross-cultural situations and to relate appropriately in a variety of cultural contexts (Bennett & Bennett 2004: 149; see also Byram 1997, Corbett 2003, Moran 2001 and Samovar & Porter 1993).

Bolten (2000: 68) and Vatter and Zapf (2012) came up with three dimensions of intercultural competence: **Cognitive, affective and the behavior-oriented dimension**. For the affective dimension, Bolten lists that people should be able to tolerate ambiguity and frustration, acceptance and respect towards other cultures, self-confidence, empathy and the willingness to intercultural learning. The cognitive dimension is characterized by knowledge about other cultures, which entails understanding attitudes, skills, knowledge and behavior, cultural differences, as well as aspects of intercultural communication and critically distancing of own communication. The third dimension is characterized through willingness and readiness to communicate, social competence and consequence of behavior, which means the ability to turn attitudes into actions (communicatively as well as non-communicatively). Erll and Gymnich (2021) have a similar definition of intercultural dimensions, as they define the affective and the cognitive dimension in the same way, but they call the third dimension **pragmatic/ communicative dimension** (Erll & Gymnich 2021: 11). However, the content of that third domain is very similar in Erll and Gymnich's model in comparison with Bolten's model.

Byram and Fleming (1998: 9) state that a person who has established intercultural competence "has knowledge of one, or, preferably, more cultures and social identities and has the capacity to discover and relate to new people from other contexts for which they have not been prepared directly" (Byram & Fleming 1998: 9). This definition is based on Byram's model of intercultural communicative competence, which will be looked at in greater detail in the next section. Fantini (2000) lists similar aspects for intercultural competence or successful intercultural communication: knowledge, skills, attitudes, awareness and language proficiency (Fantini 2000: 28). Both definitions expanding the intercultural competence with a communicative aspect. Deardorff came up with the so-called "pyramid model" of intercultural competence (Deardorff 2004: 196) and turned it into a process-oriented model (Deardorff 2006: 256). The basis is that the individual possesses attitudes like respect, openness and curiosity and discovery. The second aspect is that the individual acquires knowledge and comprehension (cultural self-awareness, deep cultural knowledge, sociolinguistic awareness) as well as skills (to listen, observe & evaluate; to analyze, interpret & relate). This leads to the internal outcome

for the individual – the third factor in the model – and means that the individual accomplishes an informed frame of reference shift (adaptability, flexibility, ethnorelative view, empathy). If the individual interacts with an interlocutor or with different people, the external outcome (4) comes to play, which means that people communicate and behave effectively and appropriately in intercultural situations. This again has an impact on their attitudes and the process starts over again. The external outcome is the result of this process, but it is also affected by the first (attitudes) and second factor (knowledge & comprehension and skills) in the process. Deardorff's model (2006) provides a new perspective on intercultural competence as she looks at the process and interaction and not only at the static individual. Her model also bases on the main aspects of intercultural competence like attitudes, knowledge and skills with a special focus on the outcome and the behavior in intercultural situations. For her studies, Deardorff conducted a survey and asked scholars who researched intercultural competence what – to them – components of intercultural competence are. All of them agreed that understanding others' world /views is important for intercultural competence. "Cultural self-awareness and capacity for self-assessment, adaptability and adjustment to new cultural environment, skills to listen and observe, general openness toward intercultural learning and to people from other cultures as well as ability to adapt to varying intercultural communication and learning styles" (Deardorff 2006: 250) are factors which 19 out of 20 scholars rated as important when it comes to intercultural competence. There are more factors like for example curiosity, mindfulness and flexibility, which are also on the list, but the ones mentioned before had the highest agreement rates.

In establishing intercultural competence, there are always three steps: (1) observation, (2) analysis and (3) explanation and action. In intercultural encounters, the person first needs to observe the interlocutor's behavior without judging it. In a second step, the person can analyze the interlocutor's behavior and can come up with possible explanations for a certain behavior the interlocutor acted out based on the observer's cultural orientation knowledge. In a last step, the person can react to the behavior based on their own analyses and explanations and the intercultural communication process is successful. Intercultural communication is not successful if a person's reaction to a certain behavior

is irritation. This usually leads to negative judgement and stereotyping and generalizing, which results in avoidance and communication fails and contact breaks off.

As we have seen, communication plays an important role in intercultural encounters and when establishing intercultural competence. In order to highlight the linguistic or communicative aspect of intercultural competence, some researchers even speak of **intercultural communicative competence** (ICC) and have established models of ICC (see Byram 1997, ICC in Balboni 2006). Balboni (2006) claims that intercultural communicative competence is constructed through “verbal codes, non-verbal codes and basic cultural values” (Balboni 2006: 16). Verbal codes for Balboni are sounds, choice of words, grammar problems, text structure and socio-pragmatic aspects, whereas non-verbal codes include expressions, gestures, smells, body noises, distance between bodies, objects, clothes and status symbols (Balboni 2006: 16). As for basic values, Balboni lists space/time, public / private, hierarchy status, respect, family, loyalty and fair play. These elements come into play in communicative events such as cocktails and dinners, telephone calls, group work, speaking in public, negotiating etc. (Balboni 2006: 16). A more comprehensible and for the educational and EFL classroom context more suitable model of ICC was established by Michael Byram.

2.1.5.2.3. BYRAM’S MODEL OF INTERCULTURAL COMMUNICATIVE COMPETENCE

Michael Byram established a very wide-spread **model of intercultural communicative competence** in his famous book *Teaching and Assessing Intercultural Communicative Competence*, which he first published in 1997, then revised in 2003 and published a second edition in 2021. His model is also based on knowledge, skills and behavior (Byram 1997: 34). He is combining the linguistic definitions of intercultural competence (see Schugk 2004) with an anthropological approach (see Lüsebrink 2016) and puts it into the framework of competences, which are characterized by education science and didactics and used in EFL teaching. This holistic approach and didactic viewpoint make Byram’s model an important source for EFL teaching and consequently, his definition also underlies the CEFR and the

German education standards. Before looking at the combined ICC model, it is important to look at Byram's definition of intercultural competence.

As for intercultural communication, Byram lists five factors (Byram 1997: 34):

	Skills interpret and relate (<i>savoir comprendre</i>)	
Knowledge of self and other; of interaction: individual and societal (<i>savoir être</i>)	Education political education critical cultural awareness (<i>savoir s'engager</i>)	Attitudes relativising self valuing other (<i>savoir être</i>)
	Skills discover and /or interact (<i>savoir apprendre/faire</i>)	

FIGURE 15: FACTORS IN INTERCULTURAL COMMUNICATION (BYRAM 1997: 34)

Coming from a didactic background, Byram also focused on the assessment and researched to what extent intercultural competence can actually be assessed. He did not only list these five factors, he also came up with **objectives** for each of the five factors mentioned above, which will be explored in greater detail.

- 1) Byram defines **attitudes** as "curiosity and openness, readiness to suspend disbelief about other cultures and belief about one's own" (Byram 1997: 50). Objectives for this factor are that learners should take opportunities to engage with otherness, they should be interested in discovering new perspectives or phenomena, question values and cultural practices in the own cultural environment, adapt to a different culture while being abroad and engage in verbal and non-verbal interaction or communication through accommodating to conventions and rites (Byram 1997: 50).
- 2) **Knowledge**, according to Byram's model, includes knowledge "of social groups and their products and practices in one's own and in one's interlocutor's country, and of the general processes of societal and individual interaction (Byram 1997: 51)". Objectives for the factor "knowledge" include that learners know about the history and culture of the interlocutor's country and the own country and their relationship towards each other,

know how to get in contact with interlocutors from another country and anticipate misunderstandings caused by different cultural backgrounds. In addition, learners should know about the definition of geographical space in their own and the interlocutors' country, as well as gaining knowledge about processes and institutions concerning social and everyday life as well as social interaction (Byram 1997: 51)

- 3) The third factor is **skills of interpreting and relating**, which Byram defines as "the ability to interpret a document or event from another culture, to explain it and relate it to documents from one's own" (Byram 1997: 51). This factor sets the objectives that the learners should be able to identify ethnocentric perspectives within written works or when they appear at events, identify misunderstandings in interaction and trying to explain them on the basis of the respective cultural background. Learners should be able to mediate different interpretations in order to avoid and prevent conflicts (Byram 1997: 51).
- 4) **Skills of discovery and interaction** is the fourth factor in Byram's model and it is the "ability to acquire new knowledge of a culture and cultural practices and the ability to operate knowledge, attitudes and skills under the constraints of real time communication and interaction" (Byram 1997: 52). Objectives are for example that learners should be able to understand concepts of different cultures and their significance and connotations as well as identifying similarities and differences in interaction, such as use of verbal and non-verbal elements or appropriate combination of knowledge, skills and attitudes as well as engage in interaction based on the already established cultural knowledge of the interlocutor. Learners should also be able to draw on their established knowledge skills and attitudes in the context of mediation between one's own and the interlocutor's culture. (Byram 1997: 52)
- 5) **Critical cultural awareness / political education** is another factor of intercultural competence and describes the "ability to evaluate critically and on the basis of explicit criteria perspectives, practices and products in one's own and other cultures and countries" (Byram 1997: 53). Learners should be able to elicit and interpret values either in documents or events – be they explicitly or implicitly shown or stated – and evaluate these with reference to universally valid

standards and criteria. In addition, learners should be able to take part in intercultural exchanges and engage in interaction and mediation activities as well as negotiating on the basis of their acquired knowledge, skills and attitudes.

These factors in combination make up for intercultural competence. Byram's comprehensive model of intercultural communicative competence (ICC), combines intercultural competence, which is based on the factors mentioned above, and communicative competence, including linguistic, sociolinguistic and discourse competence (Byram 1997: 73). Byram makes clear that communicative competence and intercultural competence go hand in hand and are intertwined.

In his model, he also describes conditions or circumstances for acquiring these competences. He states that there are three locations of learning: **classroom, fieldwork and independent learning**. In classroom learning, the teacher and the learner are involved in the learning process, whereas in fieldwork, the learner is much more active and involved and the teacher usually steps in the background. The learner is fully engaged and involved in the learning process and in acquiring these competences in independent learning, as they are doing the work and the teacher is usually not present. It is not only the relationship between the learner and the teacher, but also the interaction of the learner and the environment. In fieldwork or independent learning, when for example travelling to a foreign country or learning through the media, the learners are confronted with a different culture and the learning process and the acquisition of intercultural competence results from this intercultural exchange or experience. It is "learning by experience" or "learning by doing" whereas in the classroom setting, usually the teacher (or the classmates) is the person with whom the interaction takes place, which does not count as intercultural learning.

Byram's model of **ICC has three benefits**, which are crucial for EFL teaching (Byram 1997: 70). First, the model provides a "new" ideal: the intercultural speaker. Usually, ideals are very helpful in teaching as they give an orientation or set the goal the learners should achieve. The native speaker has been the ideal in language teaching for a long time and achieving native competence was the goal. With the

intercultural speaker, Byram presents an alternative, which is much more feasible and a reaction to the criticism and to some extent also rejection of the native speaker as the ideal or the model in language teaching. This issue will be further discussed in section 2.1.5.2.7.

Secondly, Byram's ICC model is well suited for EFL teaching as it defines the competences in a very comprehensible way and establishes objectives. Objectives are key in language teaching as they provide – again – orientation and a goal, but they also make assessment easier and thus enable teachers to see if the competence was achieved or not and what elements may be missing and should be focused on. This is not only important for teachers, but also for the learners themselves as to understand at which point in the learning process they are and what needs to be done in order to fully acquire intercultural communicative competence.

The third benefit is that Byram includes the locations of learning into his model and thus describes the role of the teacher and the learner and their respective involvedness in the learning process. Intercultural learning or intercultural competence is better acquired when learners take part in intercultural exchanges and make their own intercultural experiences. It is the application of the competences in the real world that counts, based on knowledge, skills and attitudes which could be established or learned in the classroom.

2.1.5.2.4. INTERCULTURAL COMPETENCE IN THE CEFR

Intercultural communicative competence is not explicitly stated in the CEFR (2001) as a separate competence. In the chapter about language learners and language use (chapter 4) it is stated that when acquiring a foreign language, “[t]he language learner becomes **plurilingual** and develops **interculturality**. The linguistic and cultural competences in respect of each language are modified by knowledge of the other and contribute to inter-cultural awareness, skills and know-how” (CEFR 2001: 43). This means that with this foreign language, communicative as well as intercultural knowledge,

skills and know-how are being established and built upon the general competences which were already established.

It is important to differentiate between multilingualism and plurilingualism. Plurilingualism is defined in the CEFR as “dynamic and developing linguistic repertoire of an individual user/learner”, whereas multilingualism is “the coexistence of different languages at the social or individual level” (CEFR 2020: 30).

Plurilingualism is the ability to use more than one language – and accordingly sees languages from the standpoint of speakers and learners. Multilingualism, on the other hand, refers to the presence of several languages in a given geographical area, regardless of those who speak them. In other words, the presence of two or more languages in an area does not necessarily imply that people in that area can use several of them; some use only one (COE 2016: 20).

The learner “does not keep these languages and cultures in strictly separated mental compartments, but rather builds up a communicative competence to which all knowledge and experience of language contributes and in which languages interrelate and interact” (CEFR 2001: 4-5). Plurilingualism also draws on the concept of intercomprehension and builds up language learning competence. The terms multilingualism and plurilingualism seem to be separate, however, with the growing overlap between the individual and the collective sphere, Ehrhart speaks of “multiplurilingualism” (Ehrhart 2019:27), combining the two terms and concepts. The same counts for multiculturalism and pluriculturalism, where languages can be substituted with cultures.

The term intercultural awareness is mentioned in context with general competences and declarative knowledge: “Knowledge, awareness and understanding of the relation (similarities and distinctive differences) between the ‘world of origin’ and the ‘world of the target community’ produce an intercultural awareness” (CEFR 2001: 103). **Plurilingual and pluricultural** competences are mentioned in chapter 6 of the CEFR “language learning and teaching”, but not really elaborated upon. Even though there is no chapter on intercultural competence, it is clear that intercultural competence is always thought of together with communicative competence. Also, mediation is one aspect, which plays into

the field of intercultural competence, but is listed as one aspect of communicative competence as communicative activity. Even though the 2001 version of the CEFR speaks of intercultural and pluricultural competence, there are no descriptors defined.

In the 2018 and later the 2020 companion version of the CEFR, **plurilingual and pluricultural competence** can be found as addition and as another competence besides communicative competence and signing competence and these concepts are covered in a whole chapter. Chapter 4 of the 2020 companion version deals with plurilingual and pluricultural competence. There were also descriptors for plurilingual and pluricultural competence established – for the first time in the long CEFR history. This means that intercultural and pluricultural aspects only really found their way into the European framework in 2018 or 2020 and is therefore still quite new.

Plurilingual and pluricultural competence comprises building on pluricultural repertoire, plurilingual comprehension and building on plurilingual repertoire (CEFR 2018: 157).

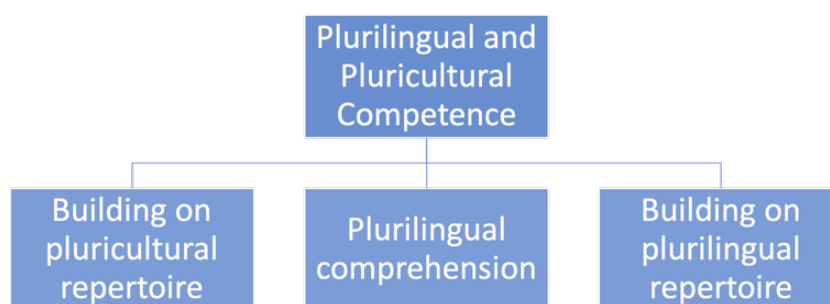


FIGURE 16: PLURILINGUAL AND PLURICULTURAL COMPETENCE

The learner is seen as a social agent and needs plurilingual and pluricultural competences in order to act appropriately in intercultural contexts and situations. I will only briefly touch upon plurilingual comprehension and building on plurilingual repertoire as the focus in this chapter is on pluricultural competence and building pluricultural repertoire.

Plurilingual competence evolves when learners establish or make use of their **plurilingual repertoire** in order to

- “switch from one language or dialect (or variety) to another;
 - express oneself in one language (or dialect, or variety) and understand a person speaking another;
 - call upon the knowledge of a number of languages (or dialects, or varieties) to make sense of a text;
 - recognise words from a common international store in a new guise;
 - mediate between individuals with no common language (or dialect, or variety), even with only a slight knowledge oneself;
 - bring the whole of one’s linguistic equipment into play, experimenting with alternative forms of expression;
 - exploit paralinguistics (mime, gesture, facial expression, etc.)”
- (CEFR 2020: 30).

Code-switching, intercomprehension, mediation and expressing and interpreting paralinguistic behavior thus play an important role in becoming plurilingual. These factors all have to do with communicative competence and come into play when learners act out communicative activities and strategies, e.g. in interaction or mediation.

Plurilingual Comprehension means knowing and making use of one or more languages in order to successfully communicate or understand texts (CEFR 2020: 126).

Key factors for plurilingual comprehension according to the CEFR are:

- “openness and flexibility to work with different elements from different languages;
 - exploiting cues;
 - exploiting similarities, recognising ‘false friends’ (from B1 up);
 - exploiting parallel sources in different languages (from B1 up);
 - collating information from all available sources (in different languages)”
- (CEFR 2020: 124).

The learners build a common (or plurilingual) repertoire based on knowledge of different languages and learn new languages with the help of languages or words they already know. General language learning competence and intercomprehension come into play here as different languages are compared to each other and used in order to make sense of texts or information.

Plurilingualism goes “hand in hand” (CEFR 2020: 30) with pluricultural competence and it is hard to only build up one competence as language and culture are always intertwined. Key concepts in **building up pluricultural repertoire** mentioned in the CEFR are:

- “recognising and acting on cultural, socio-pragmatic and socio-linguistic conventions/cues;
- recognising and interpreting similarities and differences in perspectives, practices, events;
- evaluating neutrally and critically”
(CEFR 2020: 30).

In mentioning socio-pragmatic and socio-linguistic conventions, it becomes clear that these factors also draw on communicative competence. Cultural conventions, practices and events as cultural standards or domains need to be first observed and then related and compared to other cultures, which is comparative way of interpreting culture. The first two factors are mainly concerned with the communicative aspect in Byram’s ICC model, whereas the next factor “pluricultural repertoire” draws on the intercultural aspect in the model.

Building **pluricultural repertoire** is described as

- “the need to deal with ambiguity when faced with cultural diversity, adjusting reactions, modifying language, etc.;
- the need for understanding that different cultures may have different practices and norms, and that actions may be perceived differently by people belonging to other cultures;
- the need to take into consideration differences in behaviours (including gestures, tones and attitudes), discussing over-generalisations and stereotypes;
- the need to recognise similarities and use them as a basis to improve communication;
- the will to show sensitivity to differences;
- readiness to offer and ask for clarification: anticipating possible risks of misunderstanding”
(CEFR 2020: 124).

These concept are quite in line with Byram’s ICC model and the five skills which should be established – knowledge, skills (interpretive and evaluative), education and attitude. It is also close to Lüsebrink’s definition of intercultural competence and intercultural learning as mentioned above. Thus, the factors which the CEFR lists are quite up to date. However, these are more objectives and should be further developed in order to implement them into class.

This gap was closed to some extent as in the new 2020 version of the CEFR (there were no descriptors for pluricultural competence in earlier versions) descriptors for the different proficiency levels were established (CEFR 2020: 125). In table 3 is a selection of descriptors – levels A2, B2 and C2 – for building pluricultural repertoire:

TABLE 5: DESCRIPTORS FOR BUILDING UP PLURICULTURAL REPERTOIRE

C2	Can initiate and control their actions and forms of expression according to context, showing awareness of cultural differences and making subtle adjustments in order to prevent and/or repair misunderstandings and cultural incidents.
B2	**⁶Can describe and evaluate the viewpoints and practices of their own and other social groups, showing awareness of the implicit values on which judgments and prejudices are frequently based. **Can explain their interpretation of the cultural assumptions, preconceptions, stereotypes and prejudice of their own community and of other communities that they are familiar with. **Can interpret and explain a document or event from another culture and relate it to documents or event from their own culture(s) and/or from cultures with which they are familiar. **Can discuss the objectivity and balance of information and opinions expressed in the media about their own and other communities. Can identify and reflect on similarities and differences in culturally determined behavioural patterns (e.g. gestures and speech volume or, for sign languages, sign size) and discuss their significance in order to negotiate mutual understanding. Can, in an intercultural encounter, recognise that what one normally takes for granted in a particular situation is not necessarily shared by others, and can react and express themselves appropriately. Can generally interpret cultural cues appropriately in the culture concerned. Can reflect on and explain particular ways of communicating in their own and other cultures, and the risks of misunderstanding they generate.
A2	Can recognise and apply basic cultural conventions associated with everyday social exchanges (e.g. different greetings, rituals). Can act appropriately in everyday greetings, farewells and expressions of thanks and apology, although they have difficulty coping with any departure from the routine. Can recognise that their behaviour in an everyday transaction may convey a message different from the one they intend, and can try to explain this simply. Can recognise when difficulties occur in interaction with members of other cultures, even though they may not be sure how to behave in the situation.

⁶ The asterisks (**) symbolize an advanced B-level, i.e. these descriptors could also belong to C-level (CEFR 2020: 125)

Level C2 clearly refers to cultural awareness and puts the learner in the center of intercultural situations. The learner established knowledge, skills, attitude and education through level A and B and can act culturally sensitive and appropriate in order to prevent misunderstandings. Critical incidents are mentioned as expletory situations of misunderstandings, which is a common “method” in intercultural communication to practice culturally sensitive behavior and establish skills and strategies to deal appropriately with these situations.

Level B mostly affects skills and attitude as well as to some extent knowledge and education in Byram’s ICC model. Knowledge is established and build upon further from level A, especially when it comes to norms, values and stereotypes. Relating unknown aspects of a different culture (“other culture”) to the own culture broadens their (open) attitude and understanding of cultural differences. Contrasting cultures and self-positioning within cultures are important in this level. Communicative aspects also come into play here when learners compare or are aware of different ways of communication in different cultures and to what extent this knowledge contributes to preventing misunderstandings.

Level A forms the basis of intercultural competence and is mainly concerned with knowledge and skills. The learners acquire knowledge about cultural and social rituals (greetings etc.) of their own and other cultures. They are at least aware of cultural differences, if they arise in intercultural encounters, even though they may not have strategies yet to overcome misunderstandings.

In sum, the learner establishes intercultural competence in advancing through these levels, starting from very basic knowledge about different cultures, over to contrasting the “known” with the “unknown” or “other” and establishing an awareness of differences and similarities. Acting appropriately in intercultural encounters is the most advanced level and the learners make use of communicative strategies which involve drawing on their previously established knowledge, skills, education and attitudes in order to handle these situations and prevent misunderstandings.

The COE admits that even in the 2020 version of the CEFR “[n]either pluriculturalism nor the notion of intercultural competence [...] is highly developed”(CEFR 2020:30). With their *Guide for the*

development and implementation of curricula for plurilingual and intercultural education (COE 2016)

the COE make up for this deficit with a more hands-on approach in terms of intercultural education. This guide is intended to give an orientation on how to design a curriculum for plurilingual and intercultural education and is thus aimed at curriculum designers in the European member states. It lists several aspects on how to generally design plurilingual and pluricultural curricula without listing descriptors for intercultural competence. As curriculum design is not in the scope of this study, the guide will only be mentioned here very briefly without looking at it in detail. The guide (COE 2016) also provides also a very elaborated section on mediation and key factors of mediation. What is interesting here, is that they focus on **intercultural mediation** among other aspects. The following seven competences are mentioned, which are important in intercultural mediation activities:

- “Ability to manage linguistic and cultural communication in a context of otherness;
 - Ability to construct and expand a pluralistic and linguistic and cultural repertoire;
 - The ability to move outside oneself;
 - The ability to give meaning to unfamiliar linguistic and/or cultural elements;
 - The ability to distance oneself;
 - The capacity for critical analysis of the situation and of the (communicative and/or learning) activities in which one is engaged;
 - The ability to recognize the other, recognition of otherness.”
- (COE 2016: 58).

These competences are similar to the objectives of pluricultural competence in the 2020 version of the CEFR and are probably one of many sources for their establishment.

Another important resource in the context of plurilingual and pluricultural education with a very much practical note is the *Framework of reference for pluralistic approaches to languages and cultures* (FREPA/CARAP). Pluralistic approach in this context means that more than one language and culture are used in combination in teaching, which is contrary to “singular” approaches, which focus on only one language and culture. There are four pluralistic approaches, which the FREPA mentions:

- Awakening to languages;
- Intercomprehension between related languages;
- Intercultural approach;
- Integrated didactic approach to different languages studied.

The FREPA itself defines competences and provides resources for these approaches and is designed for teachers, teacher trainers and educational leaders. As far as competences are concerned, which could be deployed with pluralistic approaches, the FREPA lists three “resources”: knowledge, skills and attitudes. These resources are the core factors when it comes to intercultural competence (see section 2.1.5.2.2.) Due to its didactic design, the FREPA also lists descriptors for each of the resources. The main advantage is that the FREPA also provides teaching material. There is the *FREPA online teaching material database*⁷, which is a collection of teaching material ready to be used in the classroom. The database can be searched for descriptors, languages, competences and topics. The topics or learning material include information about the publication year, a short description what the goal of the activity is and a description on how the lessons or the activities could be conducted. Also, the descriptors for the resources, which the activity draws on, is listed and worksheets or supplementary material is provided as the following example shows.

Additionally, there is the *FREPA – Training kit*, which can either be used for self-study or in teacher training sessions – either online or in a training setting on-site.

⁷ <https://carap.ecml.at/Materials/tabid/2313/language/de-DE/Default.aspx> (Accessed 14 March 2024).

A healthy diet



Publication year: 2011

The objectives of this activity are to make students familiar with the importance of (healthy) food and to communicate about food in the different languages they know, noticing similarities and differences between them.

Themes:

bilingual/plurilingual competence, food, iconic documents

Languages: English

★★★★★ (0 ratings, 0 comments)

[View details](#)

A 2.1	K 2.2	S 3.1.1
A 2.4.1	K 5.1	S 3.10.4
A 4.2.2	K 8.2	S 7.6.1
A 4.3.3	K 12.1.1	

Description

The objectives of this activity are to make students familiar with the importance of (healthy) food and to communicate about food in the different languages they know, noticing similarities between them. The activity is appropriate for students learning English as a foreign language or in bilingual education.

The activity is divided into two sections.

1 . Using Worksheet 1, the class discusses the issue of healthy food. Each student is asked to draw a chart with terms related to food. The words shall be translated into one's L1, a language spoken in the classroom and another language one knows. Afterwards, the class is divided into groups. Each group creates a poster about the dish of one continent using and explaining words in the respective language.

2 . During the second section, the class is asked to read a text about diet and to keep a diary about what they eat during a week. The students write down and evaluate what they ate in English or their L1. They present the results in any language they want to, as long as everybody can understand it.

Duration: 5x45

FIGURE 17: EXAMPLE FROM FREPA DATABASE (SCREENSHOT FREPA)

This screenshot from the FREPA database⁸ “A healthy diet” displays how the exercises are structured. They contain a learning objective as well as the competences (see orange, green and blue boxes), followed by a description on how to use the example in the classroom as an activity for the students.

⁸ <https://carap.ecml.at/Materials/tabid/2313/language/de-DE/Default.aspx> (Accessed 14 March 2023)

It also gives information about how long the task should be. This is of course only a suggestion; the teachers could as well adapt the activity to their learners' needs and to their own teaching plan.

Last but not least, a new publication appeared in 2022 and is called *Enriching 21st century language education – the CEFR Companion volume in practice* (COE 2022). In this document, the COE presents projects and case studies, which are based on the CEFR descriptors. The chapters in the first part for example focus on classroom practice and entail case studies conducted with regard to mediation, plurilingualism, literature and online interaction. Classroom in this context refer to secondary schools as well as classrooms in higher education. One study for example is about the implementation of descriptors in higher education, which was conducted at several universities. In the second part, the chapters focus on raising awareness and teacher education. One example is a chapter on a toolkit for language teacher training on mediation (chapter 13), which includes conceptualizing mediation, exposure to and analysis of mediation (also mediation strategies), guide to design mediation activities and reflecting on how to assess mediation (COE 2022: 203 f)

2.1.5.2.5. INTERCULTURAL COMPETENCE IN THE GERMAN EDUCATION STANDARDS

The KMK differentiates in the education standards between functional communicative competence and intercultural communicative competence and thus puts a clear emphasis on intercultural competence. Intercultural communicative competence is one of five competences which the learners should acquire when learning a language (see figure 11).

Intercultural communicative competence is defined in the German education standards as “understanding and acting in contexts where a foreign language is used. The process of understanding and acting relies on the interplay **of knowledge, attitude and awareness**” (“Wissen, Einstellungen und Bewusstheit”) (KMK 2012: 19).

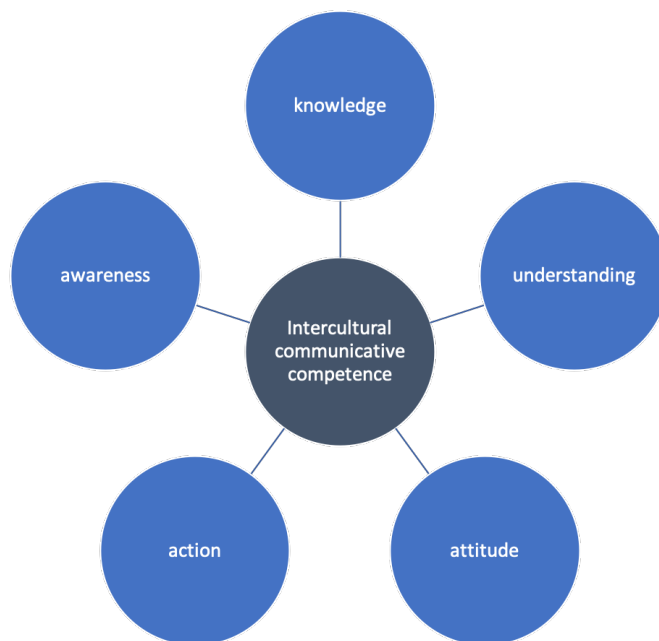


FIGURE 18: INTERCULTURAL COMMUNICATIVE COMPETENCE (KMK 2012: 19)

The over-arching goal is that the students “can communicate in direct and in media-mediated intercultural situations” (KMK 2012: 19). This goal clearly describes the two elements of intercultural communicative competence comprising of a communicative and an intercultural aspect. It is also important to note here that the mode of communication does not only include direct or face-to face interactions, but also media-mediated situations. This means that interactions via different media or digital tools, such as written or spoken online interaction in blogs or in video chats are entailed in this definition.

Intercultural communicative competence unfolds in three dimensions: **knowledge, attitude and awareness**. The learners should make use of their **knowledge** on foreign cultures in order to understand and act. They should have “socio-cultural orientation knowledge” and knowledge about cultural influences on language and language use. Strategic knowledge is also part of this dimension and it means that the learners should know a set of communication strategies in order to ensure communication when using a foreign language (KMK 2012: 19).

As far as **attitude** is concerned, the students should be able “to treat others with respect, to deal with them critically and to take risks with regard to language and content in their own language activities” (KMK 2012: 19).

In terms of **awareness**, the students should be able “to question their personal behavior and actions, enduring opinions which are different from their own point of view and reflecting in intercultural encounters” (KMK 2020: 19).

The definition of intercultural communicative competence in the German education standards is very close to Michael Byram’s model of intercultural communicative competence (Byram 1997, 2021), which is one of the most wide-spread and used models in literature for EFL teaching.

The KMK lists objectives what the learners should be able to do in the education standards:

„Die Schülerinnen und Schüler können

- ihr **Orientierungswissen** über die Zielkulturen in vielfältigen Situationen anwenden: Aspekte der Alltagskultur und Berufswelt, Themen und Probleme junger Erwachsener, gegenwärtige politische und soziale Bedingungen, historische und kulturelle Entwicklungen einschließlich literarischer Aspekte sowie Themen von globaler Bedeutung
- ihr Wissen über Kommunikation anwenden und **fremdsprachige Konventionen beachten**, u. a. zur Signalisierung von Distanz und Nähe
- ihre Wahrnehmungen und **(Vor-)Urteile** erkennen, hinterfragen, relativieren und ggf. revidieren
- einen **Perspektivenwechsel** vollziehen sowie verschiedene Perspektiven vergleichen und abwägen
- **Werte, Haltungen und Einstellungen** ihrer zielsprachigen Kommunikationspartner erkennen und unter Berücksichtigung des fremdkulturellen Hintergrundes einordnen
- fremdsprachige Texte und Diskurse in ihrer **fremdkulturellen Dimension erfassen, deuten und bewerten**
- fremde und eigene Werte, Haltungen und Einstellungen im Hinblick auf international gültige Konventionen (z. B. die Menschenrechte) **einordnen**
- ihr **strategisches Wissen** nutzen, um Missverständnisse und sprachlich-kulturell bedingte Konfliktsituationen zu erkennen und zu klären
- sich trotz des Wissens um die eigenen begrenzten kommunikativen Mittel auf **interkulturelle Kommunikationssituationen einlassen** und ihr **eigenes sprachliches Verhalten in seiner Wirkung reflektieren und bewerten**
- auch in für sie **interkulturell herausfordernden Situationen reflektiert** agieren, indem sie sprachlich und kulturell Fremdes auf den jeweiligen Hintergrund beziehen und sich konstruktiv-kritisch damit auseinandersetzen“ (KMK 2012: 20; my highlighting).

The list contains 10 objectives the learners should be able to do or fulfill (KMK 2012: 20, translated with Deep-L):

- Applying **socio-cultural orientation knowledge** of the target culture in various situations;
- Applying **foreign language conventions** (signaling distance and closeness);
- Recognize, question and relativize **perceptions and prejudices**;
- **Change of perspectives**;
- Classifying **values and attitudes** of a communication partner while taking their cultural background into account;
- Understanding, interpreting and evaluating foreign language texts in their cultural dimensions;
- Categorize **own and “other” values** and attitudes in terms of international conventions;
- **Strategic knowledge** in order to prevent (and solve) communicative and cultural misunderstandings;
- **Engaging in intercultural situations** (even with limited linguistic repertoire) and reflecting own linguistic/communicative behavior in these situations;
- Acting in a reflective manner in challenging intercultural situations so that **linguistic and cultural “otherness”** is based on cultural background and dealt with constructively and critically.

These objectives are close to the new descriptors in the CEFR and similar to Byram’s objectives in his definition of ICC. The majority of the federal states also include these objectives also into their English curricula, but unfortunately not all have yet done so (see section 2.1.5.2.6.). There are two main aspects which are still missing in the objectives: critical cultural awareness in terms of political education and *savoir s’engager*. Including critical incidents as learning examples in order to prevent cultural and communicative misunderstandings would also be helpful and should be integrated in teaching.

2.1.5.2.6. LOOKING AT GERMAN FEDERAL STATE CURRICULA

Section 2.1.4. focused on frameworks and standards in teaching at the European and the German level and in particular what the CEFR and the German Education Standards prescribe to be taught. Intercultural communicative competence plays an important role in the CEFR and the education standards and will probably even play a bigger role in teaching in the future. Communication is one of the future skills, and due to the ongoing globalization, there will be more and more intercultural encounters and collaboration in the business world, where intercultural communicative competence

is needed. As the analysis of TaCoCASE in chapter 4 and the applications (see sections 5.1. and 5.2.) show, TaCoCASE serves as a source for teaching material in the context of intercultural communicative competence. This chapter is concerned with German federal state curricula and investigates what role intercultural communicative competence plays in each of the federal state curricula.

2.1.5.2.6.1. INTERCULTURAL COMMUNICATIVE COMPETENCE AND FEDERAL STATE CURRICULA

Intercultural communicative competence is one of five competences in English language teaching according to the German education standards. While in early years in language teaching, ICC is mainly concerned with cultural traditions and norms of the target culture, in the “Oberstufe”- the introduction (“Einführungsphase”) and the main phase (“Hauptphase”) at the “Gymnasium” – intercultural communicative competence fully unfolds in many different aspects (see section 2.1.5.2.3.). The main phase of the “Oberstufe” at Gymnasium extends over the last two years before the “Abitur”. In some federal states, “Oberstufe” entails grade 11 and 12 or 12 and 13, depending on whether students spend eight (G8 system) or nine (G9 system) years at the Gymnasium in order to reach the “Abitur” exam. Due to the federal system of education in Germany, every federal state has their own federal state curriculum for each subject and it is very complicated to compare them with each other. That is the reason why the selection criteria were set to the main phase of the “Oberstufe” a Gymnasium so that these parameters were stable and a comparison between the federal state curricula is possible. Thus, the English curricula of the main phase of the “Oberstufe” at Gymnasium of the 16 federal states were taken and analyzed with regard to the role of intercultural communicative competence in these curricula.

The German “**Bildungsserver**”⁹, a website which provides links and information on the federal state curricula, was used as a starting point to collect and download the 16 curricula for each federal state.

⁹ <https://www.bildungsserver.de/lehrplaene-400-de.html> (Accessed 14 March 2023)

The government of the federal states and usually the ministry for education is responsible for publication and the composition of the federal state curricula, based on what the KMK decided and prescribed in the education standards (see section 2.1.4.2. ; figure 10).

Most of the federal states follow the KMK recommendation and list intercultural communicative competence as one of five competences in English language teaching and devote a separate chapter to the subject. This is already a good sign that ICC plays an important role in the curriculum. However, there are some federal state curricula, where intercultural communicative competence is not emphasized much and where more work needs to be done to establish it in full demand.

The KMK defines the goal of ICC in English teaching as understanding and acting in situations where a foreign language is used and that these processes are based on knowledge ("Wissen"), attitudes ("Einstellungen") and awareness ("Bewusstheit") (KMK 2012: 19). The KMK also list 10 competence expectations, i.e. what the students should be able to do or what can be expected from them (KMK 2012: 19f):

„Die Schülerinnen und Schüler können

- ihr **Orientierungswissen** über die Zielkulturen in vielfältigen Situationen anwenden: Aspekte der Alltagskultur und Berufswelt, Themen und Probleme junger Erwachsener, gegenwärtige politische und soziale Bedingungen, historische und kulturelle Entwicklungen einschließlich literarischer Aspekte sowie Themen von globaler Bedeutung
- ihr Wissen über Kommunikation anwenden und **fremdsprachige Konventionen beachten**, u. a. zur Signalisierung von Distanz und Nähe
- ihre Wahrnehmungen und **(Vor-)Urteile** erkennen, hinterfragen, relativieren und ggf. revidieren
- einen **Perspektivenwechsel** vollziehen sowie verschiedene Perspektiven vergleichen und abwägen
- **Werte, Haltungen und Einstellungen** ihrer zielsprachigen Kommunikationspartner erkennen und unter Berücksichtigung des fremdkulturellen Hintergrundes einordnen
- fremdsprachige Texte und Diskurse in ihrer **fremdkulturellen Dimension erfassen, deuten und bewerten**
- fremde und eigene Werte, Haltungen und Einstellungen im Hinblick auf international gültige Konventionen (z. B. die Menschenrechte) **einordnen**
- ihr **strategisches Wissen** nutzen, um Missverständnisse und sprachlich-kulturell bedingte Konfliktsituationen zu erkennen und zu klären
- sich trotz des Wissens um die eigenen begrenzten kommunikativen Mittel auf **interkulturelle Kommunikationssituationen einlassen** und ihr **eigenes sprachliches Verhalten in seiner Wirkung reflektieren und bewerten**

- auch in für sie **interkulturell herausfordernden Situationen reflektiert** agieren, indem sie sprachlich und kulturell Fremdes auf den jeweiligen Hintergrund beziehen und sich konstruktiv-kritisch damit auseinandersetzen.“

In order to reach a better understanding of to what degree ICC is included in the curricula, the following table (table 6) displays to what extent these expectations mentioned above are included in the federal state curricula. Column three shows that the KMK ICC competence expectation is included in the curriculum and column four shows that ICC is either not included at all or otherwise included, i.e. not meeting the ICC competence expectations. It is important to note that the latter does not mean that this is bad, ICC and the competence expectations were just defined in a different way. This was the case for Baden-Württemberg and for Nordrhein-Westfalen.

TABLE 6: ICC IN FEDERAL STATE CURRICULA

Federal State	Publication Date	KMK ICC expectation included	KMK ICC expectation not (-) or otherwise included
Baden-Württemberg	2016		x
Bayern	N/A (online)	x	
Berlin	2022	x	
Brandenburg	2022	x	
Bremen	2015	x	
Hamburg	2022	x	
Hessen	2022	x	
Mecklenburg-Vorpommern	2019	x	
Niedersachsen	2017	x	
Nordrhein-Westfalen	2023		x
Rheinland-Pfalz	1998 / 2012		x (-)
Saarland	2023	x	
Sachsen	2022		x (-)
Sachsen-Anhalt	2022	x	
Schleswig-Holstein	2014	x	
Thüringen	2019		x (-)

Table 6 shows, that the majority of curricula (11 out of 16) took the KMK ICC competence expectations and included them (almost) 1:1 into their curriculum. These federal states are Bayern, Berlin, Brandenburg, Bremen, Hamburg, Hessen, Mecklenburg-Vorpommern, Niedersachsen, Saarland, Sachsen-Anhalt and Schleswig-Holstein. Among these, ICC is usually also defined in the way the KMK defined it. Only five federal state governments did not include ICC at all or they included it otherwise than the KMK ICC expectation. Baden-Württemberg and Nordrhein-Westfalen are quite close to the definition the KMK uses, whereas Rheinland-Pfalz, Sachsen and Thüringen differ very much from the KMK definition.

In the **Baden-Württemberg** curriculum, most of the major elements of ICC are included, like communication strategies, change of perspective and socio-cultural orientational knowledge. Even though these terms are not specifically mentioned and the descriptions differ in terms of words, the general expectations and the definition of ICC is similar to the one from the KMK. ICC is also listed among the expectations in the chapter of functional communicative competence (see expectation 5 for example) as competence which is built upon. Socio-cultural orientational knowledge and topics is a separate chapter next to ICC and functional communicative competence and focuses on Great Britain and USA as target cultures with topics including globalization

In **Nordrhein-Westfalen (NRW)**, the definition and the underlying concept of the ICC is the same as in the education standards. However, the NRW curriculum does not list the exact 10 expectations from the KMK, even though they are quite similar content-wise. NRW divides the expectations into three aspects: (1) socio-cultural orientational knowledge, (2) intercultural attitudes and awareness and (3) intercultural actions and understanding with further descriptors for each category, which are similar to the expectations in the KMK. They also add specifications and suggest topics like “individual and society in transition”, “media and literature in transition” as well as “business /work and the world in transition”.

The **Rheinland-Pfalz** curriculum is a little bit more complex. While all other English curricula of the federal states have one curriculum file, the original curriculum in Rheinland Pfalz, which dates from 1998, is still in force and is supplemented by two other sources or files named on the website: Guidelines for implementing the curriculum (from 2009) and the KMK education standards (from 2012). This means that the Rheinland-Pfalz government bases English teaching on a curriculum which is over 20 years old and they have not integrated the KMK education standards, unlike every other federal state and regardless that the KMK proposed to include their standards. ICC in the KMK definition is therefore not part of the curriculum itself, but at least they refer to the KMK education standards as one of the three documents on the website. However, intercultural communication skills (“Interkulturelle Kommunikationsfähigkeit” p. 7) is mentioned and the main aspects like understanding, acting, awareness, action and knowledge are described in this section without mentioning the terms themselves. It is interesting to see that negotiation of meaning and intercultural encounters (like exchanges) also play a role in this definition. They also include “Landeskunde” as part of their concept of ICC, not in a traditional sense, but rather in a cultural studies way where acting and understanding is as important or even more important than pure knowledge. Nevertheless, some work needs to be done in order to update this curriculum in terms of state of the art ICC.

Sachsen does not mention ICC in their curriculum, but use the term “Interkulturelle Handlungsfähigkeit”, which translates as the ability to act in intercultural situations. This is only one aspect of the ICC definition in the KMK, as the KMK always uses “understanding and acting” in intercultural situations. In terms of competence expectations, the Sachsen curriculum lists only two aspects (in comparison with 10 from the KMK): The students can (1) relate or compare the own and the unknown culture on the basis of socio-cultural knowledge and (2) act as a cultural and linguistic mediator in order to prevent cultural misunderstandings. This understanding of ICC is quite rudimentary and is not very much elaborated. Another incidence of ICC can be found in LB 3 (Text and topics), where the students should have knowledge on Intercultural studies, including for example the terms “interculturality”, “globalization”, “cultural diversity” and “language and cultural identity”. In

light of the fact that the KMK published the education standards in 2012 and the curriculum in Sachsen dates from 2022, it is quite remarkable that the KMK ICC definition is not included. Therefore, I would suggest to expand the definition and the competence expectations with regard to ICC and thus put a stronger focus on ICC.

In the **Thüringen** curriculum, ICC is not mentioned with a separate chapter but the term “intercultural competence” is mentioned at the beginning in the introduction together with other basic competences. Intercultural competence is defined as an interdisciplinary competence and the description draws on socio-cultural orientational knowledge, cultural differences and similarities as well as the ability to act in multilingual situations. Some communicative strategies which are part of the KMK ICC definition are included in the chapter about mediation, but not very much elaborated. Therefore, like the Sachsen curriculum, the suggestion would be to put a stronger focus on ICC and integrating a separate chapter with a clear definition, goal and competence expectations.

Sachsen-Anhalt belongs to the federal states which included the KMK definition and competence expectations of ICC into their curriculum. However, they go even further and add some new aspects, which are quite interesting. The definition of openness and awareness with regard to other cultures for example is extended and not only entails culture but also gender and age. They also draw on sustainability and the “one world” concept in the ICC chapter, which is quite progressive. Also the use of media and in particular comparing the use of media between peers in the target countries is subject to discussion in the ICC chapter. These new aspects fit very well to the definition of ICC and extend the view and the integrability of ICC.

One of the most elaborated chapters of ICC can be found in the **Saarland** curriculum. An ICC definition, a goal and the KMK competence expectations are mentioned and three annotations are added, which go beyond the curriculum and the standard and make the implementation into the classroom more comprehensible. This is why the next section takes a closer look on the Saarland syllabus as a best practice example for ICC.

To conclude this section, we have seen that ICC plays an important role in the federal state curricula and follows in most of the cases the KMK definition in the education standards. However, there are some exceptions. On the one hand, there are curricula where ICC is defined in a different way, but still entails the major points of the concept. On the other hand, there are curricula, which find a different definition or different terms and there are some examples which put not so much focus on ICC. Especially the latter cases are definitely not standard and need to be improved in order to keep up with the fact that ICC is already playing an important role in English teaching and will even be more important in the future. In terms of comparability between the federal states and the alignment with education standards, these curricula should be revised with regard to ICC.

We can suggest that what is written in the curriculum is also part of teaching in the EFL classroom. But in fact, we do not know exactly what is going on in the classroom. From looking at textbooks and talking to teachers, ICC is mostly a competence that is taught along with functional communicative competence or other competences. Since there is no section on the Abitur which is about ICC, it is usually not assessed. And it is very difficult to access in fact. But if it is not assessed, it is usually also not taught or not taught in too much detail as the backwash assumes. However, I would argue that ICC is very important in terms of future skills and due to its rising importance in the business world it needs to be focused on in even more detail. We need to integrate more intercultural examples and teach elements of functional communicative competence like vocabulary or grammar based on intercultural examples. Why not turning this focus around? The implementation of critical incidents through either authentic encounters such as the data provided here or virtual exchanges could be a good opportunity to enhance ICC at school. And this also means, that at some point we should assess ICC as a competence and look at how students are able to prevent misunderstandings and successfully and effectively communicate in intercultural situations or encounters. With the decision of assessing ICC, the next step would be to establish a grading scheme for assessing ICC and make it part of the Abitur. This would be something which goes beyond the scope of this work, but would be very interesting and worthy of further research in. Michael Byram's suggestions in his famous work *Teaching and Assessing*

Intercultural communicative Competence (Byram 1997; 2021) could be a good starting point as objectives and the vision of the intercultural speaker is described in detail along with suggestions for assessing ICC.

2.1.5.2.6.2. BEST PRACTICE: INTERCULTURAL COMMUNICATIVE COMPETENCE IN THE SAARLAND CURRICULUM

The **Saarland curriculum** has been updated in 2023, but the ICC part had already been part of the previous version which dates from 2015. There had been only minor adjustments, but the analysis in this section will be based on the new curriculum from 2023. As already mentioned above, the main elements of ICC which were established by the KMK are integrated in this syllabus as well. This means that it is in line with what the KMK prescribes. There are, however, new aspects and new perspectives integrated, which go beyond the KMK ICC definition and which are exemplary compared to the other federal state curricula.

First of all, the definition of ICC which is used in the Saarland curriculum differs slightly from the KMK definition. While the KMK defines ICC as “understanding and acting in contexts in which a foreign language is used” (KMK 2012: 19; my translation) and thus focuses on understanding and acting, the Saarland curriculum defines ICC as “oral and written discourse and action skills” (Saarland curriculum 2023: 34; my translation). The Saarland curriculum uses the term discourse ability instead of understanding, which in my interpretation rather stresses a linguistic or communicative aspect of ICC whereas the KMK definition includes also intercultural aspects, as understanding could be comprehension in a linguistic or communicative way, but also understanding in terms of knowing what somebody means in a cultural sense. In the Saarland curriculum, the mode of communication – oral and written – are additionally mentioned in contrast to the KMK definition, which stresses both ways of communication but which is also not missing in the KMK.

The other factors like knowledge, attitudes and awareness are also mentioned and further elaborated. The same is true for the competence expectations, which were 1:1 included in the Saarland curriculum

from the KMK education standards. In terms of attitude, it is interesting to see that the Saarland curriculum stresses the term critical tolerance (Saarland curriculum 2023: 34; my translation) towards own and foreign “normality” and sees critical tolerance as a positive shift of attitudes. Critical tolerance is understood as a process of integration rather than evaluating or judging. Closely related to critical tolerance is the shift of perspectives and with it comes a “three-step”: (1) observation or description, (2) (empathically) understanding or explanation and (3) integration or evaluation of different perspectives. The goal is attaining a status of critical tolerance (Saarland curriculum 2023: 34) as described above. This three-step describes the development how a shift of perspectives can be undertaken and how students can go through this process in a quite detailed and comprehensible way. At the end of this process is again the critical tolerance, which had been defined before as one of the main goals in terms of attitudes (see Byram 2021).

So far, the focus was on theory, which includes the definition of ICC and the goal of ICC in EFL teaching with competence expectations. The best part about the Saarland curriculum is the integration of notes on to what extent the aspects mentioned before can be integrated into the classroom.

The Saarland curriculum therefore lists

- notes for methods or occasions for exemplary work
 - notes for the design of learning tasks with a focus on ICC
 - notes for assessing ICC
- (Saarland curriculum 2023: 35-36; my translation).

In implementing these notes, the chapter of the Saarland curriculum goes far beyond the curriculum border and is very helpful as a resource for teachers on how ICC can be taught in the EFL classroom. This is very unique and makes this curriculum truly an example of best practice.

Methods and occasions for exemplary work in ICC teaching

Exemplary work is very important in ICC teaching and the Saarland curriculum owes this fact a special note in the chapter of ICC. These include (Saarland curriculum 2023: 35; translated with DeepL):

- using authentic texts (audio, audio-visual and readings) in order to make intercultural potential visible

- analysis of culture-specific discourse with examples from audio texts or videos, including non-verbal facial expressions, gestures, intonation, speech rhythm, pauses, volume, eye contact, proximity and distance behavior, turn-taking, (in)directness, politeness conventions)
- analysis of cultural specific vocabulary (associations, idioms, connotations, metaphors and hot words) with the help of text, audio and video files
- mediation tasks, in which on own and foreign cultural influences are made aware
- contrastive analysis of successful intercultural communication and critical incidents with the help of excerpts from (literary) texts or audio and video material
- testing possibilities in order to make problematic communication situations a success
- critical discussion of culturally divergent positions towards a common benchmark (like UN human rights convention, etc.) with the goal of integration rather than evaluation or judgement (critical tolerance)

It is interesting to see that this exemplary work needs authentic examples and situations where cultures meet. Linguistic competence, including lexical, socio-linguistic and pragmatic competence, is drawn upon as well as interaction and mediation strategies. In looking at chapter 5 and the previous sub-chapters of this chapter, it becomes clear that TaCoCASE or ViMELF are a rich source for intercultural examples to be used in a teaching context, and excerpts from these corpora can serve as examples for all of the cases mentioned above.

Designing learning tasks with a focus on ICC

Usually, teaching in “Oberstufe” is not guided by textbooks unlike in earlier years of EFL teaching. Teachers often select texts or audio/video examples and design their own teaching material. Therefore, it is helpful to have some ideas so that what is taught in the classroom meets the demand of the curriculum.

These notes in the Saarland curriculum refer to designing tasks with a focus on ICC (translated from Saarland curriculum 2023: 36; translated with DeepL):

- selection of topics in which culture-specific perspectives meet or become clear (e.g. attitude towards weapons, speed limit, availability of medicines, organic products, data surveillance, death penalty, religiosity, immigration and integration, humor, patriotism, drugs and addictive substances);
- contrasting individual fates, groups of people, which represent different perspectives, but counteracting unreflected stereotyping;

- using supportive measures for dealing systematically and consciously with what is depicted in particular through the “three step” process of shift in perspective (describe – explain – evaluate) in the sense of critical tolerance;
- demanding a conscious shift of perspective and integration of multiple perspectives;
- productive exercises are usually simulations. Using digital media and methods allow real intercultural encounters, e.g. E-Mail projects, eTwinning, Erasmus+ projects, participation in chats and blogs which should be (self-)critically analyzed and reflected upon.

All these elements are important aspects when it comes to ICC teaching and they should be considered very carefully when designing a task for learners. They are easy to follow and comprehensible so that using them in the classroom is possible. In addition to the topics of discussion mentioned under the first bullet point, I would argue to include more feasible topics such as food, customs and traditions, sports or popular culture. These topics are closer to the students real world experience and their everyday life and don't need too much research or going into detail in contrast to data surveillance for example. I would suggest to start with more feasible topics with examples taken from TaCoCASE and ViMELF for example and then move on to “deeper” and more severe topics as mentioned in the curriculum. Another advantage of using examples from TaCoCASE and ViMELF in this context is that these are not simulations or made up; they are authentic examples from intercultural encounters. They could in a first step be analyzed with regard to interaction and mediation strategies. The students could then in a second step apply these strategies in productive activities such as simulations or role plays in class or even in digital intercultural project as mentioned in the curriculum.

Assessing ICC

This is quite a short note stating that ICC is always connected to other competences and can therefore not assessed in isolation. It usually appears together with tasks to acquire communicative competence, for example in a mediation task. It is true that ICC usually goes hand in hand with communicative competence or other competences, but the focus in assessment is usually on these other competences and not ICC. I think ICC will become even more important in the future and that is why we should also try to separate it with regard to assessment. If it is not assessed, it will not be focused on sufficiently

in teaching (backwash effect) and ICC is too important to be left aside and only taught alongside other competences.

2.1.5.2.7. THE INTERCULTURAL SPEAKER AS THE (NEW) MODEL IN EFL TEACHING?

After having discussed communicative and intercultural (communicative) competence, it is important to draw the light to a still ongoing debate in foreign language teaching, that is the debate about the aim or model in EFL teaching. The question in the center of the debate is: Is the native speaker still a suitable model for EFL teaching or are there new models which are more suitable? This section looks at different models and concludes with my own suggestion on what model to use.

The **native speaker** has for a long time been the model in language teaching. Especially in teaching methodologies like the audio-lingual method (see section 2.1.3.3.), the aim was to sound like a native speaker, use vocabulary like a native speaker and apply grammar rules like a native speaker without making any mistakes. The native speaker had a “mystical status” (see Mukerhjee 2002) and the status of proficiency or mastery of the language stems from the fact that the native speaker is born (or raised) with English as a first language, the so-called “birthright mentality” (Piller 2001). This led to “decontextualized” vocab and grammar teaching (Wilkinson 2020: 284) and a primary focus on lexical and grammatical accuracy in order to “communicate fluently and correctly, and ideally ‘pass’ as native speakers” (Wilkinson 2020: 285). This concept of the native speaker had been questioned during the last decades by many scholars (see Byram 1989, Cook 1999, Paikeday 1985, Piller 2001, Philipps & Gonzalez 2004). There are three main problems this concept of the native speaker itself and with putting the native speaker as the model for EFL teaching. The first problem has to do with the definition of the native speaker and the “birthright mentality” (Piller 2001) it suggests. The narrow definition of native speakers includes only first language speakers, i.e. they learned English from birth. These would be speakers from the inner circle in Kachru’s model (see section 2.1.3.1.). What about speakers who have more than one native language? Speakers of English as a second language for example would

also not fall into this narrow categorization and varieties of English would thus not be taught. According to this definition, people who moved to a country and spent a great amount of time in their new “home” country could never reach native speaker status. This brings citizenship and national identity in combination with linguistic competence, which is the second problem. Cook (1999) brings this definition to absurdum in stating that learners “could never become native speakers without being reborn” (Cook 1999: 187). This means that this ideal of a native speaker is unattainable and in a way frustrating for language learners as they could never achieve native speaker status and thus never full mastery of a language. But what does full mastery mean? This question leads to the third problem: perfectionism. Native speakers supposedly speak their native language perfectly and learners are seen as “deficient language users” (Kramer et al. 2014). Many scholars argue that for learners, it is not necessary to speak perfectly or 100% correctly, but it is more important to communicate effectively (Byram 1997; 2008; Cook 1999; Philipps & Gonzalez 2004; Piątkowska 2015). This means that getting a message across is more important than following all grammar rules and using vocabulary appropriately. The ultimate goal in effective communication is mutual understanding, which could involve non-standard forms. Effective communication is also the goal in ELF situations. Therefore, many researchers argue that integrating ELF language and ELF strategies in language learning is very important.

The **ideal native speaker** as the model in EFL teaching is therefore a concept, which is outdated and “useless” (Piller 2001: 121) and Paikeday (1985) even states that “[t]he native speaker is dead”. Cook (1999) argues that “language teaching would benefit by paying attention to the L2 user rather than concentrating primarily on the native speaker” (Cook 1999: 185). However, corpus linguists who look at spoken data often include native speakers as speakers in their corpus. They, do not follow this idealistic point of view or concept as McCarthy found that “[m]any of the native speakers in a corpus will be less proficient than many non-native speakers known to us” (McCarthy 2001: 140). This draws on the definition of native and non-native in Kachru’s model, but also takes into account real world experience and authentic language data. Native speakers thus sometimes also break grammar rules

and do not always follow norms. This is a very modern linguist's point of view, looking at actual language in use in contrast to the ideologized mentalist schools who favor the *ideal native speaker* concept.

If we are aware of the fact that native speakers are not “perfect” and we also integrate varieties of English into teaching, this concept of an *authentic* native speaker would be an attainable goal for language learners and could be a suitable model. The fact that effective communication is more important than perfectionism leads to the concept of “near native competence” or “native-like usage”. This includes the modernist image of the native speaker looking at actual language in use and thus becomes an attainable goal and a suitable model for learners (Mukherjee 2002).

Near-native competence and native-like usage still focus only on communicative competence, including linguistic, sociolinguistic and pragmatic competence. Cook (1999) suggests we view learners as “multicompetent language users” as they draw on many competences when speaking with other native speakers or other L2 speakers. Different competences come into play when L2 learners talk to L2 learners (Baker 2012; Risager 2006), usually also including different cultural backgrounds.

Thus, if we take it seriously that language and culture need to be taught in combination and learners should acquire communicative and intercultural competence, the ideal of the **intercultural speaker** seems very feasible (Byram 1991; Risager 2006; Roberts et al. 2001). This concept combines intercultural competence with a communicative aspect. Looking solely at intercultural competence does not work either. Intercultural competence could also unfold, if people talk in their L1 when being confronted with other cultures. In many intercultural encounters, English is used as lingua franca so that people with different first languages and different cultural backgrounds can still communicate with each other. In the language classroom and especially in the EFL classroom, learners acquire a new language and thus need to establish communicative competence. Therefore, the communicative aspect is undoubtedly important. Byram (2008) used the term “intercultural mediator”, stressing the mediation competence of the intercultural speaker in facilitating communication. Other concepts

highlighting intercultural competence include “intercultural being” (Philipps & Gonzalez 2004) taking on a humanist perspective and seeing the individual as a human being engaging in intercultural situations. When talking about transnationalism and border lessness, the “cosmopolitan speaker” as introduced by Ros i Solé (2013) is usually taken as a model. According to this concept, cultures are fluid and not stable and the cosmopolitan speaker has “multiple cultural alliances” and develops a “nomadic borderless lifestyle” (Ros i Solé 2013: 327). Another similar concept is critical cosmopolitanism where “cultural threads” are performed, in which similarities and connections are foregrounded (Holliday 2020). These concepts do not really account for a learner context as the model has to fit to the classroom context. Learners are learners of a language inside the classroom and language users outside the classroom (see section 2.1.3.1.). These concepts only unfold in intercultural situations, which the learners are usually prepared for in the EFL classroom, but these situations do not take place in a classroom setting – except in school exchanges or close enough in simulations. Thus, these concepts could only account for language users or learners in terms of lifelong education outside the classroom. Additionally, the language or communicative aspect is missing in these concepts even though undoubtedly important in the EFL classroom.

It is the combination of communicative and intercultural competence that makes the concept of the **intercultural speaker** a suitable model. Language and culture should be taught in combination and leads to the term “*linguaculture*” (Risager 2006) or “*linguaculture*” (Risager 2006; Fantini & Tirmizi 2006). But what does one qualify as intercultural speaker? Kramsch (1998b) sees the intercultural speaker as acting in a “third space”, the space between language and culture (Kramsch 1998b; see also Jones 2002). Intercultural speakers negotiate meanings and concepts and “operat[e] at the border between several languages or language varieties, mov[e] his/her way through the troubled waters of cross cultural misunderstandings” (Kramsch 1998b: 27). Mediation and opening new perspectives is according to Coperias Aguilar (2002: 92) a key asset of the intercultural speaker. Fantini (2000) list the following attributes for an intercultural speaker: “respect, empathy, flexibility, patience, interest,

curiosity, openness, motivation, a sense of humor, tolerance for ambiguity, and a willingness to suspend judgment” (Fantini 2000: 28). For Byram, an intercultural speaker is

[s]omeone who can operate their linguistic competence and their sociolinguistic awareness of the relationship between language and the context in which it is used, in order to manage interaction across cultural boundaries, to anticipate misunderstandings caused by differences in values, meanings and beliefs, and thirdly to cope with the affective as well as cognitive demands of engagements with otherness (Byram 1995: 25).

There is a kind of anthropological approach as the intercultural speaker recognizes cultural differences (Byram 1997) and would like to learn more about “the unfamiliar” or the “unknown” (Risager 2006), which drives their motivation. Cultural exchanges in field work outside the classroom or integrating cultural experiences in the classroom by cooking “foreign” recipes supports this anthropological way of learning and establishing intercultural competence (Byram 1991; 2021). Visits to other countries or virtual or real exchanges cross the classroom border and make establishing an intercultural communicative competence possible (Fantini 2019; O’Dowd & Lewis 2016). This contributes to “education for mobility”, which especially EFL teaching should enforce (Byram 2021: 91).

Byram defines what an intercultural speaker is by providing descriptions for each objective in his ICC model. The description what an intercultural speaker is able to do in combination with the objectives for the attitude dimension are displayed in the following table (Byram 1997: 57; 2021: 84):

TABLE 7: INTERCULTURAL SPEAKER AND ICC OBJECTIVES (BYRAM 1997:57; 2021: 84)

Attitudes	
Objectives	The intercultural speaker is...
(a)willingness to seek out or take up opportunities to engage with otherness in a relationship of equality, distinct from seeking out the exotic or the profitable.	The intercultural speaker is interested in the other's experience of daily life in contexts not usually presented to outsiders through the media nor used to develop a commercial relationship with outsiders; is interested in the daily experience of a range of social groups within a society and not only that represented in the dominant culture.
(b) interest in discovering other perspectives on interpretation of familiar and unfamiliar phenomena both in one's own and in other cultures and cultural practices	The intercultural speaker does not assume that familiar phenomena — cultural practices or products common to themselves and the other — are understood in the same way, or that unfamiliar phenomena can only be understood by assimilating them to their own cultural phenomena; and is aware that they need to discover

	the other person's understanding of these, and of phenomena in their own culture which are not familiar to the other person.
(c) willingness to question the values and presuppositions in cultural practices and products in one's own environment	The intercultural speaker actively seeks the other's perspectives and evaluations of phenomena in the intercultural speaker's environment which are taken for granted, and takes up the other's perspectives in order to contrast and compare with the dominant evaluations in their own society.
(d) readiness to experience the different stages of adaptation to and in interaction with another culture during a period of residence	The intercultural speaker is able to cope with their own different kinds of experience of otherness (e.g. enthusiasm, withdrawal) during residence and place them in a longer term context of phases of acceptance and rejection.
(e) readiness to engage with the conventions and rites of verbal and non-verbal communication and interaction	The intercultural speaker notes and adopts the behaviours specific to a social group in a way which they and the members of that group consider to be appropriate for an outsider; the intercultural speaker takes into consideration the expectations the others may have about appropriate behaviour from foreigners.

In a similar manner, he also lists descriptions in combination with the objectives for the other three dimensions: skills of discovery and interaction, skills to interpret and relate, knowledge and critical cultural awareness / political education (see Byram 2021: 85 ff).

Byram established this concept of the intercultural speaker already in 1997 and it is obvious that the CEFR and also the education standards draw on his model of ICC without really stating that the intercultural speaker is the new aim or model. The CEFR for example states that “[t]he language learner becomes *plurilingual* and develops *interculturality*” (CEFR 2001: 43; my highlighting). This would clearly fit to the concept of an intercultural speaker. The native speaker is not the model any more by stating that “the aim of language education is profoundly modified. It is no longer seen as simply to achieve ‘mastery’ of one or two, or even three languages, each taken in isolation, with the ‘ideal native speaker’ as the ultimate model” (CEFR 2001: 5). However, they are reluctant to explicitly state that the intercultural speaker is the new model, which according to their definition would qualify as the better option. Also the aim defined in the education standards in Germany saying that “students are able to

act and communicate in direct or mediated intercultural situations” (KMK 2012: 19, my translation) implicitly draws on the concept of the intercultural speaker.

In recent years, Byram came up with an even broader concept that of *intercultural citizenship*. For him, intercultural citizenship education involves two aspects:

- Causing/facilitating intercultural citizenship experience, and analysis and reflection on it and on the possibility of further social and/or political activity, i.e. activity that involves working with others to achieve an agreed end;
- Creating learning/change in the individual: cognitive, attitudinal, behavioural change; change in self-perception; change in relationships with others [i.e. people of different social groups]; change that is based in the particular but is related to the universal (Byram 2008: 187).

This concept is based on the intercultural competence model but adds “action” as a fourth dimension, besides “attitudes and feelings”, “behaviour” and “knowledge and skills” (Byram 2012). According to Byram (2012), taking action involves to speak about intercultural encounters and thus informing other people about intercultural topics, raising intercultural awareness and sensitivity. He even goes further and describes a more complex form of taking action by being involved in civil society organizations (Byram 2012). In my opinion, this model can only be distinguished from the intercultural speaker through “action” which is a very personal and also political decision. Therefore, I do not consider this a useful aim for ELF teaching and would prefer the intercultural speaker in this case .

To sum up, this section presented some concepts which are around the debate of the aim or model in ELF teaching. It is clear that the ideal native speaker as a model is outdated and not very well suited in a modern EFL classroom. The concept of authentic native speakers as usually used in corpus linguistics and related to that the near native competence as an orientation seems to qualify for establishing communicative competence. Especially in early years of learning where much emphasis is put on establishing a linguistic repertoire and communicative competence, this goal seems fitting. However, as Byram (2012: 89) puts it

[t]he intercultural speaker needs intercultural communicative competence, i.e. both intercultural competence and linguist/communicative competence in any task of mediation

where two distinct linguacultures are present and this is something different from and not comparable with the competence of the native speaker (Byram 2012: 89).

Intercultural communicative competence is becoming more and more important in ELF teaching and the model should move from the native speaker to an intercultural speaker. As I am trying to encourage people to take a broader view in this work by looking at the goal of ELF teaching and preparing learners for the real world, in which communication and collaboration in digital and global environments is needed, it is important to establish intercultural communicative competence. And by arguing to establish intercultural communicative competence, it seems only reasonable to argue for an intercultural speaker as the goal in ELF teaching.

2.1.6. EFL TEACHING AND LEARNING IN THE DIGITAL ERA

Migration and globalization are not the only challenges we are facing in today's world. Another main challenge is digitalization and the digital transformation. Education plays a key role in turning these challenges into opportunities for our society and it is the task of the education system and all people involved in education to prepare learners in the best possible way for the future, i.e. to live and work responsibly in such a globalized and digitalized world. The EFL classroom has traditionally been the place to learn how to cross communicative and cultural borders through establishing communicative and intercultural competence. In recent years, as mentioned above, communicative competence and intercultural competence have become more and more important and these competences will still be key in EFL learning. However, digital competence plays an important role in today's world and in the future, which means that learners need to be prepared for that as well. Digitization has had a revolutionary impact on education and on each subject, including EFL teaching, and that is why we need to integrate aspects of digitalization also in the field of EFL teaching. The following section looks at how digitization influenced teaching with regard to the teaching process and digital tools and what competences go with it – from a learner's as well as a teacher's perspective.

2.1.6.1. FROM DIGITALIZATION TO THE CULTURE OF DIGITALITY

Digitalization revolutionizes the way we live and the way we work. Digitization is more than using a digital device or a digital tool; digitization changes the way we communicate, the way we collaborate, the way information and data are stored, shared and made accessible. Digitization and digital education has a technological aspect and is usually centered around digital tools. However, this is only one aspect. According to experts who signed the „Dagstuhl-Erklärung“, digital education needs to be looked at from a **technological**, a **social-cultural** and an **applied perspective** (Dagstuhl-Erklärung 2016: 1). They established the well-known “Dagstuhl-Dreieck” (Engl.: Dagstuhl-triangle), which describes

that phenomena, items or objects and situations should be viewed from the following three perspectives (Dagstuhl-Erklärung 2016: 3):

- Technological perspective: How does it work?
- Social-cultural perspective: Which effect does it have?
- Applied perspective: How do I use it?

Additionally, they state that digital education should be integrated in all subjects and it is a continuing process for learners throughout the whole school curriculum. Digital education should also be included in teacher training in all phases and research in subject-specific didactics needs to be conducted and new ideas for integrating digital education into the respective subjects need to be developed (Dagstuhl-Erklärung 2016: 1)

In a similar manner, the KMK defines digitization in an educational context as follows:

Die Digitalisierung unserer Welt wird hier im weiteren Sinne verstanden als Prozess, in dem digitale Medien und digitale Werkzeuge zunehmend an die Stelle analoger Verfahren treten und diese nicht nur ablösen, sondern neue Perspektiven in allen gesellschaftlichen, wirtschaftlichen und wissenschaftlichen Bereichen erschließen, aber auch neue Fragestellungen z.B. zum Schutz der Privatsphäre mit sich bringen (KMK 2016: 8).

They stress the importance of digital tools and their use in a sense that they do not only substitute analog material, but that they provide new applications and new ways of thinking are possible. This is certainly the overall objective and describes the state at the end of the digital transformation. The way, however, towards this objective is long and at the moment, the education system is still in the middle of this process.

This process is characterized through three main steps: **Digitization**, **digital transformation** and the **culture of digitality** (see Döbeli Honegger 2021: 41-42). The first step is digitalization, including technological change and with it the change of the leading medium. With the implementation of computers, the way how data and information are stored and shared has changed fundamentally. In this step, analog processes and items were transferred into the new, digital medium. The way we can access data and work with (big) data – including computer tools and automatization – influences the

way we work and processes have started to change. Communication and collaboration have also changed when using digital, interconnected tools and programs. This leads to the second step, **digital transformation**. Analog information and processes are not only transferred into the digital medium, they begin to start changing objectives and new, innovative and disruptive products and processes evolve, which would not have been possible without digital tools or the digital medium. The last step is the **culture of digitality** (“Kultur der Digitalität”, see Stadler 2016) in a so-called post-digital perspective, which describes a state where the process of digital transformation is completed and digital tools and processes are fully implemented in our lives. Felix Stadler coined the term “Kultur der Digitalität” or culture of digitality by stating that digitality is

a set of relations, which is realized today on basis of digital infrastructure and networks of production, use and transformation of material and immaterial items, as well as the constitution and coordination of personal and collective action (Stadler 2016: 18, my translation).

Digitality, according to Stadler (2016), produces new possibilities and interconnects human and non-human actors, which does not limit the term digitality to the digital medium, but expands it to a relational pattern, which could appear everywhere (Stadler 2016: 18). He also states that in the culture of digitality, the analog does not completely disappear and it is still present, but re-evaluated and upgraded.

These phases can be transferred into the educational context with regard to digital tools. A widely used model is the SAMR-model (Puentedura 2006), which describes four different technological levels of use:

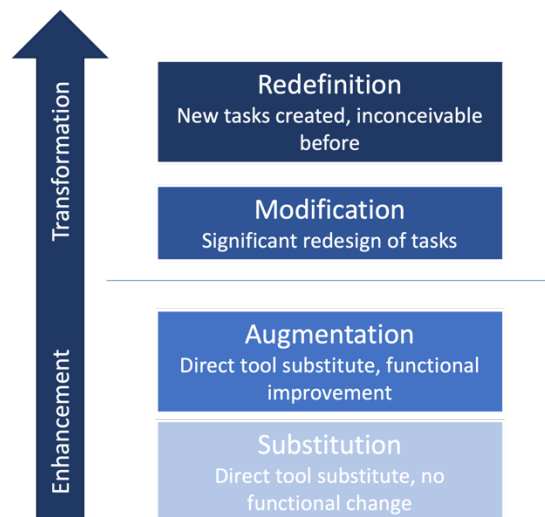


FIGURE 19: SAMR-MODEL (ADAPTED FROM PUENTEDURA 2006)

These four levels describe technological levels of use and will be explained in a classroom context (see Pfiffner et al. 2021). *Substitution* means that an analog process or material is simply transferred into a digital tool or digital material. In the education context, this would be for example a handout which was a print material before, turned into a pdf and shared with the learners. So, only the medium has changed, content and all learning processes and tasks stay the same.

Augmentation is the step after substitution and in this step, the analog tool is substituted with the digital tool and tasks stay the same. However, in this step, functions improve as there exist new possibilities how to use the tool or how to use the material. An example would for example be a QR-Code which could be scanned and which links to an online dictionary or a video, which gives provides more information on a different medium than before. Augmented reality apps are also a good way to connect the analog with the digital world. With the next step, we leave the phase of enhancement and enter the phase of transformation.

In the *Modification* step, the medium changes from analog to digital and the content or task are being redesigned. New tasks and new ways of collaboration are possible for example when learners work together on a task with different digital sources. The last step and the most advanced step of digital transformation in the SAMR model is **Redefinition**. In this step, medium, content and task are

completely digital and they would not work in the analog world at all. Examples for this step would be visualizations in digital tools, collaborative tasks which are carried out in the digital or virtual space or virtual reality tools.

These steps form the basis for technological development and integration as part of the digital transformation. I would argue that the majority of the education material and tools in the classroom are still on the Substitution and Augmentation level and therefore only enhance digital learning. More tasks should be developed in order to reach modification and redefinition and thus move forward in terms of transformation in the educational context and finally end up in digitality rather than digitization. It is important to note, that there is no right or wrong in using these tasks and there is no value in the levels in a sense that only higher tasks and tools in the SAMR model are valuable. It is still the teachers' methodological decision how to integrate these tasks and on what level the tools and tasks operate.

The process of digital transformation takes place inside and outside the classroom and the learners experience digital transformation outside the classroom very often, for example in computer games. Additionally, the working environment has also changed and is on the way of digital transformation, which calls for new skills and competences the learners need. In a logical consequence, we do not only need to drive the digital transformation further in our classrooms, we also need to prepare the learners for their future (work) life. Before looking at future skills and new competences, I would like to discuss some challenges of digital education in the classroom.

In principal, there are **three kinds of positions**: People who reject digital tools completely, people who think that digital tools and digital media need to prove their additional value and people who drive digital transformation by integrating digital tools on all levels of the SAMR model. This counts for the world outside the classroom as well as for the world inside the classroom, where the term people could be substituted by teachers. The first kind of teachers who reject digital tools simply reject reality and the ongoing digital transformation in our lives. They do not take their task in educating and

preparing students for their future seriously. Luckily, this number of teachers is very small in contrast to the other two categories as more and more teachers see the need to integrate digital tools and tasks into the classroom. The second category can be described as supporters of the **“Mehrwerttheorie”** (translates to “added value theory”). Krommer (2018) summarizes the viewpoint of supporters of this category as follows:

1. Digital tools should not only be used because they are digital and digital teaching does not lead per se to better teaching.
2. Using digital methods is additional effort in terms of time, organization, technology, finances and didactics.
3. This additional effort is only justified if the curricular objectives can be achieved faster and more easily (see Krommer 2018, my translation)

Baumgartner and Herber (2013) argue in a similar way, by stating that interactive media have didactic potential (**“Didaktisches Potenzial”**) and a didactic (added) value (**“Didaktischer Mehrwert”**) (Baumgartner & Herber 2013: 331). According to them, the didactic potential results in the combination of interactive or digital medium together with the teacher and a didactic setting. Didactic (added) value is only possible if the learning effect with interaction of the medium is “superior” to the learning effect in situations without these kinds of media (Baumgartner & Herber 2013: 331).

The **“Mehrwerttheorie”** is often criticized in a sense that analog tools and tasks do not have to prove their value, whereas digital media have, which is regarded as unfair (Krommer 2018). Another argument against the **“Mehrwerttheorie”** is that the learning effort is hard to measure. Using digital tools means that proficiency and assessment need to be adapted to these digital media. It is odd to say that “old” assessment categories can be used with “new” media, as for example collaborative skills and communicative skills cannot be tested in a multiple choice test. Additionally, learning effects need to be seen in a broader category. If for example learners use corpora and corpus tools in teaching, it takes longer to learn how to use these tools, but at the same time, they acquire competences which help them using similar tools in the future. Therefore, learning effects need to be measured in terms of new competences and with new assessment methods.

What is undisputed is the first aspect of the Mehrwerttheorie or the didactic potential as introduced above. Digital tools should not be used simply out of technological reasons, but “in a meaningful combination of digital and non-digital media” (Ladel 2018: 13, my translation). It is the combination of analog and digital media, which is sometimes a good option. And above all, the teachers and methodology plays an important role in this complex situation as “the pedagogy and not the technology drives the activity” (Goertler 2009: 75). Döbeli Honegger echoes this by stating that „science as well as practice have shown for years that it is not the digital tool or medium, but the didactically wise use at the right time, which can promote teaching and learning” (Döbeli Honegger 2021: 44). Using digital tools is therefore not a threat for teachers, but a “more” on methods and thus a broader range in terms of methodology.

2.1.6.2. 21ST CENTURY SKILLS AND DIGITAL COMPETENCES

As mentioned above, the digital transformation changed our way of life, how we communicate and how we collaborate. This calls for new skills and competences which are needed in a globalized and digital world, and which therefore should be established at school. In a first step, we will look at new skills and competences for learners and teachers before taking a look at how the German education system defines digital education and what skills and competences are suggested.

The most widely known competences in the digital era are 21st century skills (Schleicher 2013) and the 4K-model, which is used in Germany and is based on Schleicher’s model of 21st century skills.

21st century skills, according to Andreas Schleicher, who is an OECD education researcher and PISA coordinator, are the following:

1. Ways of thinking: creativity, critical thinking, problem-solving, decision-making and learning
2. Ways of working: communication & collaboration
3. Tools for working: information literacy, technology

These skills count for everyone and are not restricted to learners. The 4C-model (or in Germany 4K-Modell) is based on the 21st century skills and lists the most important ones:

1. **C**ritical thinking
2. **C**reativity
3. **C**ollaboration
4. **C**ommunication

Depending on the source, sometimes creativity is supplemented with innovation and critical thinking with problem-solving (Pfiffner et al. 2021: 29; see also Muuß-Merholz 2017, Mihajlović 2017 and Poitzmann & Sobel 2024: 14f). Creativity for example means thinking (and working and learning) in new ways (Muuß-Merholz 2017). Critical thinking comprises of problem solving and having the ability to think (work and learn) for oneself. Collaboration includes thinking (and working and learning) together with others and communication means being able to share what one is thinking (and working and learning) with others (see Muuß-Merholz 2017). These 4 C's should not be regarded as individual competences but rather in combination with each other (see Poitzmann & Sobel 2024: 15).

Another model which focuses on competences or knowledge teachers should have, is the **TPACK-model** (Koehler et al 2014: 103) (see figure 20).

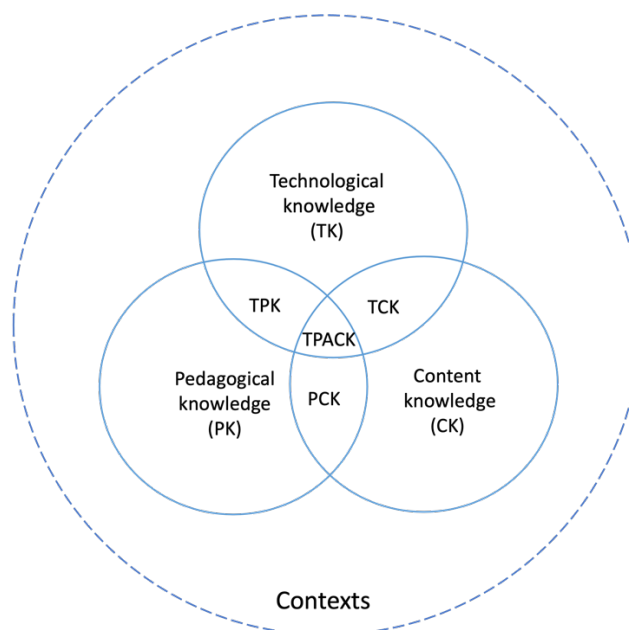


FIGURE 20: TPACK-MODEL (ADAPTED FROM KOEHLER ET AL. 2014: 103)

This model describes three kinds of knowledge teachers should have: pedagogical knowledge, content knowledge and technological knowledge. The TPACK model depicts these competences as circles overlapping each other, which indicates that they should be regarded in combination with each other. TPACK is at the heart of the circles and draws on all other circles, resulting in the combination of all three kinds of knowledge and in the term **Technological Pedagogical And Content Knowledge**. This model reminds of the didactic triangle (see section 2.1.1.), while pedagogical knowledge in this case also includes content and methodology as well as objectives. It adds technological knowledge, which is a reasonable addition in light of digital education.

The European commission established a framework for teachers as well, called **DigCompEdu** (a Digital Competence Framework for Educators). This framework differentiates between educators' professional competences, educators' pedagogic competences and learners' competences. Among professional competences, professional engagement, including for example organizational communication, and professional collaboration can be found. As for pedagogic competences, there are digital resources, assessment, teaching and learning and empowering learners with several competences for each respective category. In the area of learners' competences, the competence listed is facilitating learners' digital competence, including for example information & media literacy and content creation among others (DigCompEdu 2017: 8).

The following section focuses on Germany and the standards and competences the KMK suggests for digital teaching.

2.1.6.3. EDUCATION IN THE DIGITAL ERA IN GERMANY

There are two main strategy papers the KMK published in order to describe digital education in Germany, set competences and provide suggestions for teaching. In 2016, the KMK published a strategy paper called ***Bildung in der digitalen Welt*** for implementing digital education in teaching

which focuses not only on language teaching, but teaching in general. This strategy paper looked at education from a digital perspective for the first time and sets competence expectations for digital education. In this paper, the KMK speaks of a digital revolution in the world and in teaching. Their concept of digitization includes aspects of digital transformation as digital tools not only substitute analog tools, but they open new possibilities and perspectives not only in education but in our lives (KMK 2016: 8). They see opportunities – changes of education processes and promoting talents and individual potential – and challenges – new ways and environments of learning, new learning objectives, infrastructural and personal conditions (KMK 2016: 8). Above all stands the “Primat des Pädagogischen” (translates to primacy of the pedagogic; KMK 2016: 9), which stresses that the teachers and the methodological decisions for using digital tools is more important than the digital tool itself. Digital tools “support” teaching, but teaching and the overall education objective of preparing learners for their future life is the principal aim of teaching (KMK 2016: 10). A short-term goal was to provide access to a digital learning environment to every learner until 2021 (KMK 2016: 11), which I would argue has not been reached everywhere. This is due to the fact, that some requirements, which the KMK lists, like infrastructure, legal issues, further development of teaching and qualification of teachers have not been completely fulfilled, even though advances in every aspects have been made.

The paper suggests competences which should be established for a **general digital competence**. These should be integrated into the subject-specific curricula and adapted to the respective subject-specific methodologies, even though not all competences can be established in the same degree in all subjects.

The competences include (KMK 2016: 15; translated with Deep-L):

1. Searching, processing and storing
2. Communicating and cooperating
3. Producing and presenting
4. Protecting and acting safely
5. Solving problems and taking action
6. Analyzing and reflecting

In 2021, the KMK published additional recommendations to the strategy paper for teaching and learning in a digital world under the title ***Lehren und Lernen in der digitalen Welt – die ergänzende Empfehlung zur Strategie “Bildung in der digitalen Welt”***. In this strategy paper, they broaden the educational goal by explicitly stating that the culture of digitality is the state which is the aim. They implicitly draw on the 4K-model as they list the four competences in this model as the competences which should be established – critical thinking, creativity, communication and collaboration – and add “acting competently” as a fifth competence (KMK 2021: 8).

The paper lists the following potentials of a digital learning culture (KMK 2021: 9; translated with Deep-L):

- Individuality and creativity
- Experimentation and discovery-based learning
- Action and product orientation
- Variability and switching between forms of presentation
- Vividness and multimedia
- Extended communication and collaboration
- Connection of different learning locations (e.g. hybrid learning settings, learning cooperation)
- Support for accessibility
- Individual feedback and feedback on the learning process

Indeed, these potentials could all be reached with digital tools and provide possibilities of new ways of learning. There are several measures which lead to a better digital learning culture and to fully unfold the potentials mentioned above (KMK 2021: 12; translated with Deep-L):

- Increasing motivation and enjoyment of learning
- Simulations, dynamic modelling, collaborative problem solving
- Combining digital and analog teaching and learning settings
- Accessibility, i.e. inclusive access
- Intelligent digital learning environments for the diagnosis of learning statuses, support of teaching-learning processes and material as well as task and examination formats for individual support (in cooperation with science, if possible, freely accessible)

These are very important measures which have an effect on the learner's learning outcome. Digital tools like games or apps, increase motivation as learners are still curious about using technologies. They probably even use more digital tools and apps outside the classroom already, which definitely calls for integrating a reflective use inside the classroom as well. One of the main advantages of using digital tools is that learners could get individual feedback right away in an app for example and that their learning journey can also be individualized. With the growing differentiation, this could support learners and teachers as faster learners could obtain more and maybe also more difficult tasks and slow learners could have less tasks. Tasks, content and complexity could be adapted to the learning style and thus guarantee a better learning outcome for the individual learner.

Measuring learning outcome under these changed conditions is one major challenge for the education system. To date, many exams or assessments are still mostly written. In EFL teaching, more and more oral exams and presentations became standard. Using digital tools and methods in teaching and establishing digital competences, this also needs to be reflected on in assessments: "Future examination formats will therefore increasingly incorporate creativity, collaboration, critical thinking and communication" (KMK 2021: 13). If the 4K's will also be assessed, new forms of assessment like simulations or collaborative problem solving will become the new standard. In my opinion, the EFL classroom is very well suited for simulations and collaborative problem solving as communication skills are needed and established. Intercultural simulations for example could fit perfect to the EFL classroom and to establishing digital, communicative and intercultural competence at the same time.

Digitization and EFL federal state curricula in Germany

Since these publications and the standards count for education in general, the KMK demands that these digital aspects should be included in the curricula for the different subjects. Not all English curricula had been updated since these new digital papers had been published and thus the digital aspects are not yet included in each of the English curricula. By the way, the education standards for

English have also not been adapted according to digital education, which will hopefully be soon the case.

Some English curricula like the ones from Hamburg, Mecklenburg-Vorpommern, Sachsen-Anhalt and Saarland for example, include aspects of digital education. They all state that English as a Lingua Franca and English as the subject itself play an important role in a globalized world. And with globalization, digitization also play an important role.

- **Mecklenburg-Vorpommern** for example states that English is the leading subject for communicating using various digital methods and selecting digital communication strategies (Englisch curriculum Mecklenburg Vorpommern 2019: 5).
- **Sachsen-Anhalt** argues that English as a subject contributes to education in the digital world and that the combination of EFL teaching with digital media and tools contributes to active participation in a globalized world. With the help of digital tools, students can control and individualize their learning process. The Sachsen-Anhalt curriculum lists competence expectations based on the KMK strategy paper “Bildung in der digitalen Welt” (English curriculum Sachsen-Anhalt 2022: 12) and thus makes very clear how digital tools, methods and media can be integrated and what is expected from the students. These are the competence expectations from the Sachsen-Anhalt curriculum (English curriculum Sachsen-Anhalt 2022: 13; translated with DeepL):
 - find, document, share and securely store information,
 - critically assess different digital sources and their foreign-language content, especially with regard to their reliability,
 - reflect on media, media content and media use with regard to their possible effects,
 - interact with the help of digital communication options in the foreign language in a targeted manner and in a manner appropriate to the situation and addressee,
 - process, present and publish foreign language content using various digital tools, taking into account legal requirements,
 - operate safely in digital environments by recognizing, reflecting on and taking into account their risks and dangers,
 - develop solution strategies for dealing with technical problems and
 - assess the opportunities and risks of their own media use in different areas of life

- The **Hamburg** curriculum for example also draws on the role of digital media in the EFL classroom: “Digital media can simulate the impression of real target language environments” (English curriculum Hamburg: 2022: 10). They argue that it is important to acquire digital competences in the target language as well in order to participation in a/synchronous digital communication situations in the target language. This represents an important skill in today’s world and in the future and will also motivate the learners.
- The **Saarland** curriculum mentions that digital tools and artificial intelligence can support the improvement of functional communicative competence and communicative activities. According to the Saarland curriculum, digital media should be integrated into the English language classroom through receptive activities as well as (collaborative) production activities. They also mention participation in a/synchronous digital situations in the target language. They even go one step further and invite the students to critically reflect on digital tools in terms of “critical media literacy”. What is interesting to see is that the Saarland curriculum allows assessment of production skills (monological and dialogical) with digital media so that the students can for example invent and present podcasts, blogs and thus use computer-mediated or digital media to support the production activity. This is a very good example how digital tools can support English language teaching. Since it is close to the students’ real world experience and prepares them for their future (business) world, the students have a higher degree of motivation and it can be assumed that the learning process is lasting longer and is thus more effective.

The digital transformation influences all areas of our lives and thus also teaching. According to a recent survey, 79% of parents think that digital competences are important for the future of their children and they are open to integrating digital technology into the classroom (Vodafone Stiftung 2023: 5). Digitization and driving digital transformation is a process that needs to be actively shaped by policy and decision makers in education, teachers, parents and learners. It is a challenge that needs to be

integrated in all subjects and according to subject specific needs. The fact that some federal states updated their curricula in terms of digital competences is a good sign, but more work needs to be done. It is also important not to leave the teachers alone in a sense that teachers should be educated how to use digital tools and technology in language learning in further education and more material should be provided for hands-on use in the classroom. This is probably one of the major challenges and the task for the present and the future. One way of integrating digital tools and technology is data-driven learning and the use of corpora in language teaching. Using corpora in the EFL classroom is very well suited to establish digital competences as well as communicative and intercultural competences – as will be discussed in chapter 5. The following section explains what corpus linguistics and corpora are and to what extent the use of corpora can support EFL teaching.

2.2. CORPUS LINGUISTICS AND THE ROLE OF CORPORA IN EFL TEACHING

This sub-chapter builds the bridge between EFL teaching and corpus linguistics. While the previous section looked at EFL teaching and some general didactic and methodological issues, this section introduces corpus linguistics as a field of linguistics and explains what corpora are, what advantages they have, and how they can be used by looking at key methodological questions. The focus then will be on how corpora can be used in EFL teaching.

2.2.1. FOUNDATIONS OF CORPUS LINGUISTICS

In this section, the focus will be on corpus linguistics and corpora in general. I will answer questions like what corpora are and what they can be used for and what the research focus is in the field of corpus linguistics, before some corpus methods will be explored. This sets the groundwork for the following sections as the data used in this work is a corpus and some analyses in chapter 4 are based on corpus methods.

2.2.1.1. WHAT IS A CORPUS?

The term **corpus** itself is Latin meaning “body”. Susan Hunston, an influential corpus linguist with a focus on applied linguistics who wrote the renowned book *Corpora in Applied Linguistics*, claims that “a corpus is defined in terms of both its form and its purpose” and continues with her definition stating that a corpus, for linguists, has always been “a collection of naturally occurring examples of language, consisting of anything from a few sentences to a set of written texts or tape recordings, which have been collected for linguistic study” (Hunston 2002: 2). With the rise of computers and electronic devices in the last half of the 20th century, the definition has been updated and narrowed to “collections of texts (or parts of texts) that are stored and accessed electronically” (Hunston 2002: 2). The size of corpora, the plural of corpus, has expanded as computers can store larger samples of texts than texts collected manually. Thus, in a linguistic sense, a corpus is “a body of naturally occurring

language” (Baker 2006: 1) where the intention of collecting the texts was to find out more about **language in use** with the help of electronic devices. This is another important criterion that distinguishes a corpus from an archive or a library, i.e. that “a corpus is planned [...] and designed for some linguistic purpose” (Hunston 2002: 2). Computers cannot only be used for storage reasons, but also for analytical reasons, which means that a corpus can be studied “non-linearly, and both quantitatively and qualitatively” (Hunston 2002: 2). I will have a look at how corpora can be analyzed and with what kind of methods later in this chapter.

John Sinclair (2005) who is also a major researcher in the field of corpus linguistics, stated that “[a] corpus is a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research” (Sinclair 2005). His definition sheds light on another important aspect: a corpus describes language of a specific variety or genre depicted in the collected texts. This means, that the purpose and the fact that texts are representative of a specific genre, are important information about the corpus.

Gries (2016) defines a corpus in the same way as Hunston and Sinclair and highlights the purpose of the collection by stating that “[a] corpus refers to a machine-readable collection of (spoken or written) texts that were produced in a natural communicative setting, and in which the collection of texts is compiled with the intention (1) to be representative and balanced with respect to a particular linguistic language, variety, register, or genre and (2) to be analyzed linguistically” (Gries 2016: 7).

In sum, a corpus is a resource for analysis consisting of texts that represent authentic language and it is a collection of texts which can be analyzed with computer software in order to find out more about language in use. Corpora have “revolutionized the study of language” (Hunston 2002: 1) and the way linguists can find evidence for language in use as “[c]orpora allow researchers not only to count categories in traditional approaches to language but also to observe categories and phenomena that have not been noticed before” (Hunston 2002: 1).

Types of corpora

Depending on the purpose of the corpus, there are several **types** of corpora (see Hunston 2002: 14):

- General corpora: collection of texts of many types, large corpus; examples: British National Corpus (BNC; 100 million words) or Corpus of Contemporary American English (COCA; 560+ million words)
- Specialized corpora: specific to a certain type of text, e.g. newspaper articles, textbooks, essays written by students
- Comparable corpora: texts of two different languages or language varieties; example: International Corpus of English (ICE)
- Parallel corpora: “two (or more) corpora in different languages, each containing texts that have been translated from one language into the other” (Hunston 2002: 15)
- Learner corpora: texts produced by learners; International Corpus of Learner English (ICLE)

Another way of classifying corpora, which McEnery & Hardie (2012) mention, is distinguishing corpora according to specific features such as **mode of communication: written and spoken corpora** or monolingual and multilingual corpora. Parallel and comparable corpora would fall into the category of multilingual corpora, whereas the BNC and COCA would be examples of monolingual corpora. Written corpora consist of written texts such as novels, newspaper articles or student essays. They are relatively easy to build as there is lots of data already existent in electronic format (McEnery & Hardie 2012). In contrast, spoken corpora, which consist of spoken language like in any kind of recording or (video- or computer-mediated) interaction are harder to build. Spoken language needs to be recorded at first and in a second step it needs to be transcribed so that it can be stored as a text. There are several aspects which need to be taken into account when transcribing spoken language (see section 3.1.2.3.2.). Due to the fact that spoken language has many additional features compared to written language, like pauses, interruptions or overlap, it is very time-consuming to transcribe. These factors play an important role when compiling a spoken corpus and in fact they limit the size and the availability of spoken corpora. Spoken corpora play an important role in corpus linguistics, as Caines et al. (2016) point out:

In comparison with written corpora, spoken corpora have not developed at the same rate. The reasons for this are largely to do with the huge costs and time involved in compilation and

transcription, as well as access to recordable data. What has developed over the last 20 years, however, is an acknowledgement of the importance of spoken corpora in creating a fuller understanding of everyday spoken language, especially casual conversation. (Caines et al. 2016: 348)

Caines et al. (2016) also see the limitations of spoken corpora mentioned above and claim that spoken corpora are important to look at and should therefore be researched to a larger extent. However, this is only possible with more spoken corpora available and they see a lack of casual conversation data.

Video-mediated communication is even more difficult as there are features like non-verbal communication which carry important meaning and also need to be transcribed in some way. TaCoCASE and ViMELF, the corpora investigated in this work, are spoken corpora and contain video-mediated communication between international students. This multi-modal setup makes these corpora difficult to build, but in fact, they depict even a more authentic and deeper image of spoken language – casual conversation – in contrast to audio recordings.

Criteria for building corpora

Besides the fact that a corpus is designed for a special purpose, there are other criteria which are important when compiling a corpus or criteria to bear in mind when working with corpora, such as **size**, **content**, **balance** and **representativeness** of the corpus (see Hunston 2002). These factors come into play at once and are crucial elements which help to make decisions in corpus design and during the compilation process. These criteria are approached and explained in this section from a theoretical point of view. They form the basis for the design of TaCoCASE, which will be explained in section 3.1.2.1. more in detail.

Size is a very controversial issue and a quite visible criterion when comparing or selecting corpora. The range starts from some very large corpora, such as the BNC or COCA with more than 100 million words to smaller corpora with several thousand up to a million words like for example ELFA or VOICE. As

already mentioned, written corpora are relatively easy to collect, while the compilation of spoken corpora with audio or video material and the production of the respective transcripts is more time-consuming. Hunston pointed out that “investigation that depend on the manual annotation of the corpus are also, necessarily, restricted in terms of the size of the corpus that can feasibly be annotated” (Hunston 2002: 26). The fact that the data should be manageable for the researchers is a limiting factor in corpus compilation. Spoken corpora tend to be much smaller compared to written corpora. O’Keeffe and colleagues (2007) give a specific number or categorization of what is considered to be large and small for spoken and written corpora: “For a corpus of spoken language, anything over a million words is considered to be large; for written corpora, anything below five million is quite small” (O’Keeffe et al 2007: 4). Corpus size is also dependent on the **research focus**. Carter and McCarthy (1995: 143) for example argue that a smaller corpus is sufficient when analyzing grammar words as they tend to be very frequent in spoken language. When it comes to **content**, the question is what should be integrated in a corpus and what not. Hunston (2002) makes very clear that “decisions about what should go into a corpus are based on what the corpus is going to be used for, but also about what is available” (Hunston 2002: 26). Thus, use and application of the corpus play a key role in corpus design. When designing a corpus, the researcher should have a clear picture in mind what the corpus could be used for and also how accessible the data is that the researcher wants to collect.

Another two important elements are **balance and representativeness**. “A balanced corpus might be said to consist of equal numbers of words in each category” (Hunston 2002: 28). Whether the corpus is balanced according to number of words or hours of recordings or other factors, it does not matter so much as long as there is an equal distribution of a specific factor and a stable variable. Along with a balanced corpus there usually comes **representativeness**. According to Biber, “a corpus must be ‘representative’ in order to be appropriately used as the basis for generalizations concerning a language as a whole” (Biber 1993: 243). This means that corpora can for example represent the language of a particular speaker group like ELF speakers or EFL learners. In order to make generalizations, the corpus must be representative with regard to that specific aspect. This again

depends on the language in the corpus and what the researcher aimed at when designing the corpus. The focus of the analysis therefore plays an important role when picking a corpus or as Hunston puts it: "The real question as regards representativeness is how the balance of a corpus should be taken into account when interpreting data from that corpus" (Hunston 2002: 31).

It is important to have these criteria in mind when working with corpora as they influence decisions on what corpora to use for what kind of research. They are crucial when compiling a corpus because these criteria drive the compilation and the decision making process.

Annotation and key terms

After having looked at important criteria in the corpus compilation process, the focus now is on some key terms and key aspects of **storing the corpus data**. A corpus usually consists of text files, either written texts or spoken language transcribed and converted into text. These text files are stored electronically on a computer or in a webspace like an online document management system. The text files contain the actual text and a header, which provides additional information on the circumstances of collection and the files themselves. TaCoCASE, the corpus investigated here, for example contains information like filename, copyright information, information about the medium and the duration of the recording. In order to use computer programs and software to work with the corpus and to make up for differences in characters for many languages, the text files usually need to be saved as .txt file and Unicode UTF8 as character encoding is recommended (McEnery & Xiao 2004). Most corpus analysis tools nowadays also work with Word files, but working with txt-files is more precise as different characters are encoded and leave less room for interpretation.

Systematic mark-up of the texts includes annotation. Annotating texts is the activity of tagging and parsing a corpus. Adding a tag to a word in a corpus means adding additional information to the word. A **tag** is thus an ID code, which describes a word with its function, for example its grammatical function. The most prominent kind of tagging is **part-of-speech tagging (POS-tagging)**, which means that the

tag, which is added to a word, entails information on the part of speech of the word. Tagging usually works automatically and does not need to be done manually (any more). One very good example for an online tagger is the CLAWS POS tagger¹⁰ which was developed at Lancaster University and can be accessed and used for free. The tag *_NN1* for example means “singular common noun” (e.g. boy) according to the CLAWS 7 tagset. There are many different tags so that each individual word in the corpus can be tagged according to this tagset with its POS function.

When looking at larger units than the individual word, the annotation type is called **syntactic tagging or parsing**. Parsing is “the analysis of texts into constituents, such as clauses and groups” (Hunston 2002: 19) and it displays the number of different structures in a text. Semantic tagging means adding semantic tags to the words. Semantic tags “show semantic fields which group together word senses that are related by virtue of their being connected at some level of generality with the same mental concept” (Archer, Wilson & Rayson 2002). UCREL Semantic Analysis System (USAS)¹¹ is a good semantic tagger which was also developed at Lancaster University and can be accessed for free. In this tagset, *L* stands for example for “life and living things” and *E* for “emotion” (see Archer, Wilson & Rayson 2002 for more examples).

Spoken language is not very easy to tag as there are features like overlap or laughter, which could not easily be displayed by the tags mentioned above. The TaCoCASE and ViMELF transcripts were converted to **XML (Extensible markup language) format**, which is similar to HTML (usually used for websites). While HTML tags indicate how the content should be displayed (on a website for example), XML tags indicate how the content can be described. XTranscript or BaseX are free online tools for converting transcripts into XML format. XTranscript works very well with the CASE data as the CASE transcription conventions were integrated into the tagset (see section 3.1.2.2.).

¹⁰ <https://ucrel.lancs.ac.uk/claws/trial.html> (UCREL 2018a)

¹¹ <https://ucrel.lancs.ac.uk/usas/> (UCREL 2018b)

When analyzing corpus data, it is important to use the correct terms. Some **key terms** will be discussed very briefly. A single word in a corpus, meaning a linguistic unit which is separated by spaces or punctuation, is called **token**. Each different word (or linguistic unit) in a corpus is called **type**. If the item is repeated it is only counted once. A “lemma” is a category used for related word forms and is usually spelt in capital letters. The nouns *house*, *houses* and *housing* would all belong to the same lemma *HOUSE*. *BE* would be the lemma for the word forms *am*, *are* and *is*.

As we have now laid the foundation of what a corpus is, how corpora can be distinguished and what criteria are important for their compilation, we will now have a look at what we can do with a corpus and how we can analyze language using corpora.

2.2.1.2. WHAT IS CORPUS LINGUISTICS?

Corpora, as we have seen in the previous section, display actual language in use and can be used for various applications. **Corpus linguistics** is, according to McEnery and Hardie (2012: no pages), “a field which focuses upon a set of procedures, or methods, for studying language”. They continue and state that “[i]mportantly, the development of corpus linguistics has also spawned new theories of language – theories which draw their inspiration from attested language use and the findings drawn from it” (McEnery & Hardie 2012, no pages). Corpus linguistics has thus had a great impact on the way language was studied. Stefanowitsch (2019) claims that “[c]orpus linguistics is the investigation of linguistic research questions that have been framed in terms of the conditional distribution of linguistic phenomena in a linguistic corpus” (Stefanowitsch 2019: 56). The aim in corpus linguistics is using **corpus methods** to answer **linguistic research questions** looking at a corpus, which displays actual language in use, which Tognini Bonelli puts like this: “Given that a corpus is a collection of texts, the aim of corpus linguistics has rightly been seen as the analysis and description of language use, as realised in text(s)” (Tognini Bonelli 2010: 19). Corpus linguists use corpus methods and computer

programs in order to study language. This is one of the reasons why it is still considered still quite a new approach to language analysis.

Brief history of Corpus Linguistics

Corpus Linguistics started in the second half of the 20th century. With the appearance of the first computers and the rise of digitization, corpus linguistics and corpus methods found their way into linguistics and the study of language. Also, linguistic questions got more and more complex and called for empirical evidence. In the 1950s, “[c]ompiling a collection of texts which would provide sound evidence of the state of language was new” (Tognini Bonelli 2010: 15). Around that time, Randolph Quirk started to collect data in order to conduct grammatical research, which resulted in the *Survey of English Usage*, a collection of spoken and written data comprising of about 1 million words. Based on this work, Quirk and colleagues (1985) published *A Comprehensive Grammar of the English language* 30 years later as well as the *International Corpus of English (ICE)*. The first electronic corpus, however, was the **Brown corpus**, compiled at Brown University under Nelson Francis and Henry Kučera in the 1960s. In the early 1970s, the LOB (London-Oslo-Bergen) corpus was compiled by Stig Johansson and colleagues, following the same principles and structure as Brown.

In 1977, the **ICAME** (International Computer Archive of Modern and Medieval English) was founded by W. Nelson Francis, Geoffrey Leech, Stig Johansson, Arthur O. Sandved and Jan Svartvik, with the aim of bringing international corpus linguists together and providing an making corpora, the results of corpus analyses and material available to a larger audience. The ICAME conference is still one of the most important conference in corpus linguistics and the ICAME journal is also renowned in the field of corpus linguistics. During the 1980s and 1990s using computers and technology became more and more feasible and made life of corpus linguist easier.

With technical progress, corpus size began to increase as the texts could more easily be stored and collected. Corpus annotation and frequency analysis were also possible with the help of computers

and technology. More and more corpora were compiled and made accessible to other researchers. During the 2000s and the implementation of the world wide web, texts were even easier to access online and could be collected in corpora. Specialized corpora are on the rise and corpus size even gets larger, which makes the results even more representative and generalized. The British National Corpus (BNC) for example was one of the largest corpora with 100 million words in the early 1990s. Tagging was now automated, which made quantitative analysis even faster and less inaccurate compared to manual analysis. In the 2010s, corpora move online and can be accessed online. Data from the web is increasingly used for corpus analysis, which results in the “**Web as a Corpus**” approach. Collecting written texts and compiling written corpora becomes even more easy and commonplace, while spoken corpora are still the minority. The number of spoken corpora increased and the interest in spoken language is still on the rise, even though written corpora outnumber spoken corpora. Multimodal corpora, which include multimodal features such as gestures, gaze or other non-verbal behavior, are to date still very rare. The CoRD (Corpus Resource Database)¹² provides a good overview of the different corpora which are available. It is important to mention that corpus linguistics had a major influence on the compilation of dictionaries and grammar books. The *Longman Student Grammar of Written and Spoken English* (Biber et al. 1999) for example includes corpus data and *Collins COBUILD English dictionary* (Sinclair 1995) provides information on the use of words, such as grammatical or frequency information (Mukherjee 2002: 72).

Corpus Linguistics – discipline or method?

There is still an ongoing debate whether corpus linguistics is a discipline in its own right or whether it is a method to study language. On the one hand, corpus linguistics can be seen as a **linguistic method** analyzing and studying language (Leech 1992: 105; Aarts 2000: 7). Using corpus analysis tools and

¹² <http://www.helsinki.fi/varieng/CoRD/corpora/index.html>

corpus methods already implies that it indeed has strong methodological character. In this case, corpora are used as resources and data is being explored by formulating hypothesis and testing whether these hypotheses about language are valid or not. Approaches like this are called **corpus-based**. On the other hand, it is clear that corpus linguistics had a major and disruptive impact on the study of language and thus on linguistics. It provides empirical evidence of language in use and thus delivers new, groundbreaking information about language. For Sinclair, corpus linguistics is a “new approach” to analyzing language (Sinclair 1991: 2) and Tognini Bonelli goes even further by stating that corpus-driven linguistics is a “new discipline within linguistics” (Tognini Bonelli 2001: 177). Some researchers see the corpus itself as source or body of theory of language, which is considered a **corpus-driven approach** (Tognini Bonelli 2001: 84) and underpins the discipline character of corpus linguistics. Mukherjee (2002) also comes to the conclusion that it should be regarded as a **discipline**. He argues that the insights are so groundbreaking that it could count as a discipline in its own right (Mukherjee 2002: 53).

In my opinion, both views are right and depend on the approach. I would argue that only with corpus methods and by compiling corpora is it possible to gain new insights into actual language use. Thus, corpus linguistics lives from corpus methods and using corpora to investigate language in use. More and more corpora are available and thus researchers provide more and more evidence for linguistic questions. Since these results are so disruptive and the approach changes the way we investigate language, I would argue that it is a discipline in its own right. I would even encourage researchers to use corpora more often and integrate corpus methods to a larger extent into their own research as the results are coined by authentic language. Corpus linguistics draws on so many fields of linguistics and corpus-based research could provide interesting results in other linguistic disciplines as well. Corpus linguistics and its methods could also be used by teachers or learners and could have a revolutionary and disruptive impact on authentic language teaching too (see also Mukherjee 2002: 53).

2.2.1.3. HOW TO USE CORPORA

Corpora can be used in many ways and in many linguistic fields. In this section, some **basic corpus methods** will be explained and it will be further investigated how corpora can be analyzed with the help of corpus tools. The following list provides typical questions which could be investigated with the help of corpus methods (McCarthy, cited in Bennett 2010: 4):

- What are the most frequent words and phrases in English?
- What are the differences between spoken and written English?
- What tenses do people use most frequently?
- What prepositions follow particular verbs?
- How do people use words like *can*, *may*, and *might*?
- Which words are used in more formal situations and which are used in more informal ones?
- How often do people use idiomatic expressions?
- How many words must a learner know to participate in everyday conversation?
- How many different words do native speakers generally use in conversation?

This list already implies that corpus analyses provide information about how frequently a word or a pattern is used. Frequency plays an important role and is usually the starting point in the analysis of corpora. Analyzing corpora often follows the same procedure and underlying concept: quantitative findings draw the attention to frequent words or frequent clusters and these would be investigated further in a qualitative approach.

In this context, it is important to note, that corpora display actual language in use and when looking at corpora, we can only describe how language is used. There is no answer to the question whether something is right or wrong (and what would even be considered to be “right” or “wrong”?) – the way language is perceived in corpora is descriptive.

Corpus analysis tools

Before looking at some basic analyses, some **corpus tools** will be introduced and explained. Typically, for analysis of corpora, standard software is used which runs on a computer. This would be classified

as third generation concordancer as the generations before that could not be installed on own computers or could not handle a larger amount of data (McEnery & Hardie 2012). As for computer software, the most common concordance tools are **Wordsmith** (Scott 1999) and **AntConc** (Anthony 2023). With the help of these tools, corpora can be loaded into the software and further investigated. They are very helpful to analyze corpora of any size and provide results within seconds.

Lawrence Anthony has a large collection of different software programs¹³ he provides for free online, such as AntGram, AntWordProfiler and AntFileConverter. AntConc is the most common and probably most frequently used program as it contains the basic techniques and tasks for analyzing corpora.

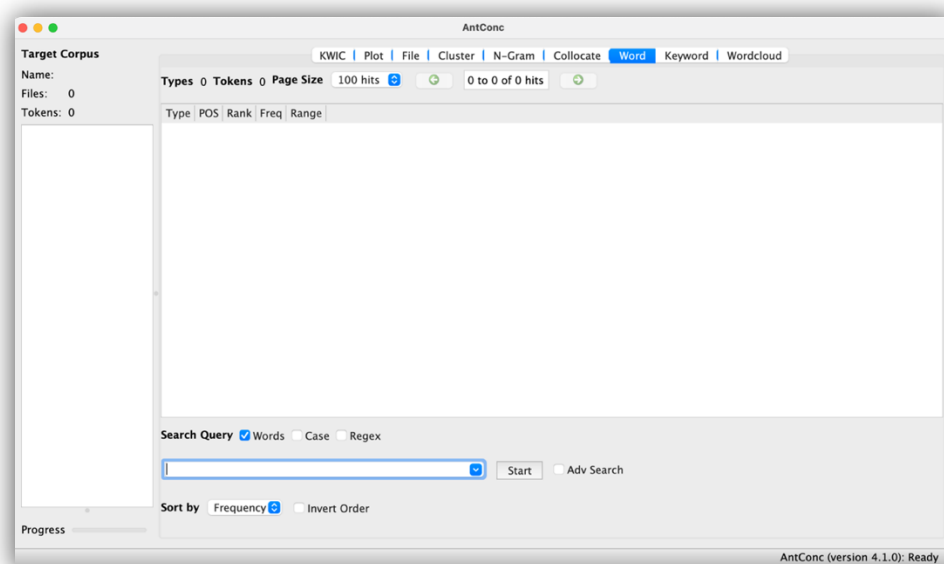


FIGURE 21: SCREENSHOT FROM ANTCONC (VERSION 4.1.0.)

Figure 21 is a screenshot from AntConc (version 4.1.0), which is very well-known and widely used by corpus linguists and also many applied linguists doing corpus-based research. The program was used on a Mac without input at all. On the left side, the corpus files can be loaded and the right window shows the results of the analysis. There are several tasks which can be conducted with AntConc, which

¹³ <https://www.laurenceanthony.net/software.html>

are listed in the upper row, like KWIC search, cluster analysis or N-gram analysis and word frequency counts. Some of these tasks or analyses will be discussed in detail below.

Nowadays, there are also **online tools** which were implemented on a website and allow some basic analysis, which is considered the fourth generation of concordancing. Some corpora even have integrated searches like COCA for example. WebCorpLSE is the tool TaCoCASE is integrated in (see section 3.1.2.2.). It is a fourth generation concordancer as it runs in a web browser and as it also has also other corpora included which could be analyzed. With these online search engines, it is important to know which corpora are integrated or used in the background for the analysis in order to make sure that the choice of corpora fits the research question. It would for example not make sense to look how hesitation markers are used in spoken language based on a written corpus.

Corpus analyzing techniques

The most common applications or analysis techniques are (see O’Keeffe et al. 2007: 11-13):

- Word frequency or word list
- Key word analysis
- Concordance analysis
- Cluster analysis

Word frequency is one of the most basic and still most widely used analyses in corpus linguistics. It is – as the name implies – concerned with frequency of words. Word frequency counts indicates how often a word appears in a corpus or in a subset of a corpus. The frequency of a specific word can then be compared to either a word with a similar function or similar meaning. The question then would be why one word is used more frequently than the other word. The frequency of a specific word could also be compared across registers, for example when comparing frequencies in a spoken and a written corpus. Words could be more frequently used in either one of the genres or there could be no difference, which would also give us information about how a specific word is used across registers.

Frequency of words in a corpus is typically displayed in a **wordlist**, which is “a rank ordering of all words in [the corpus] in order of frequency” (O’Keeffe et al. 2007: 11). With these wordlists, comparisons across corpora and across genres or registers are possible. It could, for example, be the case that in a spoken corpus, certain words are more frequent than in a written corpus or there could be only small differences. This result gives hints to what extent spoken language is different from written language – based on the corpus data and thus real language in use. One hypothesis could be that in a spoken corpus, discourse markers or hesitation markers would be more frequently used than in a written corpus. By looking at the Top 10 or Top 20 list of the most frequent words, this hypothesis could be proved valid or wrong. It is important to note here, that not only tokens can be searched for, and that it is also possible to search for lemmas or for tags, like for example POS tags.

Closely related to word frequency is **keyword analysis**. “Keywords (or key tags) [are] lists of items which are unusually frequent in the corpus or text you are investigating, in comparison to a reference corpus” (McEnery & Hardie 2012). If a word is more frequently used in a corpus or a sub-corpus compared to a reference corpus, this means that can be classified as a keyword in the respective corpus. This information is very useful in teaching English for specific purposes (ESP) like for example English for pilots (see O’Keeffe et al. 2007: 12). Typically, word or keyword frequency information gives hints about which words are worth further investigating further.

In a concordance analysis, we move away from word lists and keywords to words in context or concordances. **Concordance analysis**, or concordancing, is “using corpus software to find every occurrence of a particular word or phrase” (O’Keeffe et al. 2007: 8). Concordances are usually displayed in a concordancing tool like AntConc so that the concordances are underneath each other and the word of interest is in the middle of each concordance. The researcher can then look at the context around the word of interest or node in different concordance lines. The concordances can usually be ordered alphabetically to the left or right of the node or in terms of frequency, depending on the research focus. Concordances are usually read from top to bottom and not from left to right like a normal sentence. In doing so, the researcher is able to find patterns across concordances and

look at how a word is used in a phrase or utterance. This kind of analysis is very interesting for all kinds of research questions around phraseology.

In these corpus tools, it is not only possible to look at single words but also at combinations of words or “chunks”. **Cluster analysis** looks at frequent word combinations. Therefore, these tools provide statistical information of how frequently words co-occur. Depending on the research focus and on the length of the stretch of words, lists with frequent two or three word combinations can be generated. This means that the respective two or three words statistically occur frequently together. One example for a cluster analysis would be collocation analysis or n-gram analysis. N-grams are chunks or clusters of words which occur statistically frequent together, where n indicates the number of words in a particular chunk. A 4-gram would for example consist of four words. This analysis is very useful in EFL teaching and for vocabulary teaching in particular. Cluster analysis provides “implications for what we teach in our vocabulary lessons and how learners approach the task of acquiring vocabulary and developing frequency” (O’Keeffe et al. 2007: 13).

As we have seen, there are many analyses which could be conducted with the help of corpus tools. Depending on the research focus, either a more quantitative or qualitative approach or a combination of both can be chosen. How these analyses can be conducted in a teaching context and to what extent they can be used in the classroom and contribute to EFL teaching, will be explored in the following section.

2.2.2. USING CORPORA IN EFL TEACHING

This section combines the fields corpus linguistics and EFL teaching. The use of corpora has not only revolutionized language research, there is strong evidence that it could also have a major impact in EFL teaching if corpus linguistics was integrated more often. It will be explored how corpora and corpus methods can be used in the modern EFL classroom and what contributions they can make in EFL

teaching. We will see what influence corpus linguistics has on teaching and why it is important to integrate corpora and corpus methods in the EFL classroom.

2.2.2.1. THE AUTHENTICITY GAP

It is a fact in language teaching that many teachers use textbooks in order to guide them through teaching and be in line with the curriculum as textbooks are usually designed according to curricula. The biggest problem, however, is that language taught through textbooks – or in the classroom in general – is not always **authentic**. Morrow states that “[a]n *authentic text* is a stretch of real language, produced by a real speaker or writer for a real audience and designed to convey a real message of some sort” (Morrow 1977: 13). Yule (1995) argues that there is “a substantial mismatch between what tends to be presented to learners as classroom experiences of the target language and the actual use of that language as discourse outside the classroom” (Yule 1995: 185). This leads to demands to “bring textbooks for teaching English as a foreign language into closer correspondence with actual English” (Mindt 1996: 247).

The demand for “actual English” or authentic language use in the EFL classroom was manifested in many studies (see Fox 1987; Kennedy 1992; Römer 2005, Sinclair 1997). Sinclair (1991) states that there exists an “absurd notion that invented examples can actually represent the language better than real ones” (Sinclair 1991: 5). Kennedy echoes Sinclair and even speaks of misleading examples in textbooks by claiming that “invented examples can present a distorted version of typicality or an over-tidy picture of the system” (Kennedy 1992: 366). Using invented dialogues for teaching communicative and interaction skills has been criticized in many studies (see Carter 1998, Thornbury & Slade 2006) as they lack main spoken language features like discourse markers or hedges (see Carter 1998). Instead, it is better to use examples from authentic spoken language for enhancing linguistic competence (see Burns 2001).

Since corpora display **actual language in use** according to Morrow's criteria, they can be considered authentic. And this authenticity should also be present in the language classroom and in language learning as "[a]uthentic recordings of casual conversation are the most likely source of useful models to illustrate how proficient speakers effectively manage discourse and build relationships" (Gilmore 2007: 102). They can serve as examples for learners as to how language is used by proficient speakers. Communicative language teaching is all about integrating authentic examples and actual language in use into the classroom. Thus, "the use of authentic data is seen to avoid the risk of distorting real language patterns, illustrating how the latter are actually used to communicate in context" (Aston 2000: 12). Biber and Reppen (2002) looked at grammar teaching and grammar used in textbooks compared to uses in a corpus. Based on their studies, they suggest that "frequency should play a key role in the development of materials and in the choices that teachers make in language classrooms" (Biber & Reppen 2002: 206).

It is important to note here that many **dictionaries** and **academic grammars** are corpus based. Some dictionaries provide frequency information and the widely-known student grammar *Longman Student Grammar of written and spoken English* (Biber et al. 1999) contains examples and frequency information based on corpus analyses. Michael Barlow argues that "[t]he results of a corpus-based investigation can serve as a firm basis for both linguistic description and, on the applied side, as input for language learning" (Barlow 1996: 32). O'Keeffe et al. go even further and note that "[a] spoken corpus can underline for us how important it is to look closely at what speakers and listeners do" (O'Keeffe et al. 2007: 30) and thus "provide empirical basis for intuitions" in a sense that "language presented in textbooks is frequently still based on intuitions about how we use language, rather than actual evidence of use" (O'Keeffe et al. 2007: 21).

The use of corpora can help bridge this gap by providing **authentic text samples**. Corpora as authentic texts "provide access to naturally occurring language and corpus linguistic methods support exploratory and discovery learning" (Cheng 2010: 320; see also Bernardini 2004). McCarthy also argued for a corpus-based pedagogy in stating that "[w]ith a more accurate picture of natural

discourse, we are in a better position to evaluate the descriptions upon which we base our teaching, the teaching materials, what goes on in the classroom, and the end products of our teaching, whether in the form of spoken or written output” (McCarthy 1991: 12). He not only looked at the authentic input corpora can depict, he goes even further and states that corpus-based teaching and using corpus methods also has a positive impact on the learner’s output.

In sum, authentic texts and authentic language use as depicted in corpora, should be used in the EFL classroom. It is not only the corpus or the authentic texts though that have an impact on teaching. The following section gives insights into other influences corpus linguistics has on language teaching.

2.2.2.2. INFLUENCE OF CORPUS LINGUISTICS ON LANGUAGE TEACHING

Susan Hunston (2002) states that integrating corpus linguistics in teaching has two potential effects for teachers:

Firstly, corpora lead to new descriptions of a language, so that the content of what the language teacher is teaching is perceived to change in radical ways [...]. Secondly, corpora themselves can be exploited to produce language teaching materials, and can form the basis for new approaches to syllabus design and to methodology. (Hunston 2002: 137).

She thus states that using corpora in the EFL classroom influences *what* is taught as well as *how* it is taught. If these two fundamental elements of pedagogy – content and methodology – are influenced by the use of corpora and corpus methods, it can be concluded that corpus linguistics has a major influence on language teaching.

There are three main distinctions which can be made to what extent corpus-based linguistics can be integrated in teaching (see Aston 2000: 7, see also Fligelstone 1993, Leech 1997):

- **Teaching about corpora**, which includes teaching about the field of corpus linguistics with the aim of integrating it in university curricula – in particular in the field of linguistics – and school curricula;
- **Exploiting corpora to teach languages** and other subjects, which includes the analysis of corpora and drawing conclusions for various fields of research and other subjects
- **Teaching to exploit corpora**, so that learners can explore them on their own (corpus as resource, data-driven learning)

Sometimes a fourth dimension is added: teaching to establish resources (Renouf 1997) or having learners involved in the collection of data (Lee & Swales 2006: 70). This categorization draws on different target groups and different levels of involvedness. The first category for example is located more on a meta-level and the second and third category have a direct impact on teaching and teachers and learners respectively. However, without raising awareness for corpora in research and in teacher education at universities, it would not be integrated in classroom teaching by the teachers.

Mukherjee (2002) therefore calls for **“corpus literacy”** in a first step as depicted in the first category (see above). He argues that the way corpus linguistics can be integrated in the language classroom goes from teacher education at university and in teacher training over to the English language classroom and lastly to the learner when they are autonomously using corpus data (Mukherjee 2004: 245). As far as the EFL classroom level is concerned, he finds it “necessary to implement teacher-centred corpus activities in the classroom before truly learner-centred methods are envisaged” (Mukherjee 2004: 239).

Ute Römer (2011) distinguishes between two pedagogical corpus applications: **indirect and direct applications**. Indirect applications, according to her, are “hands on for researchers and material writers” and have an effect on both (1) the teaching syllabus as well as on (2) reference works and teaching materials (Römer 2011: 207). In this category, the target group would be people responsible for syllabus and textbook design such as researchers and publishers or textbook editors. Teachers and learners are not involved in these approaches. Direct applications, however, are tailored to hands on use by learners and teachers and are centered around data-driven learning. These kind of applications

involve “teacher-corpus interaction” as well as “learner-corpus-interaction”, which means that a corpus is being used in the classroom either by the teacher or the learner in order to facilitate language learning (Römer 2011: 207).

As the focus of this work is primarily on EFL teaching and the application of corpora in the classroom, I decided not to look at indirect applications, as in syllabus design or teaching material (even though the implications drawn in chapter 4 and 5 could be interpreted with regard to indirect applications as well). Therefore, the following sections are concerned with **direct applications** of corpora in language teaching such as data-driven learning and how corpora can actually be used in the classroom.

2.2.2.3. DATA-DRIVEN LEARNING

Using corpora in the classroom is considered to be **data-driven learning** (DDL). DDL was introduced in the 1990s primarily by Tim Johns and defined as “the use in the classroom of computer-generated concordances to get students to explore regularities of patterning in the target language, and the development of activities and exercises based on concordance output” (Johns & King 1991: iii). It is an approach in applied linguistics, a “useful learning methodology” (Cheng 2010: 320) in communicative language teaching in particular, which focuses on using actual data to facilitate language learning. John’s quite narrow definition restricting DDL to using concordances has been opened up in recent years. Today, data-driven learning includes using corpora in the classroom in general as well as “using the tools and techniques of corpus linguistics for pedagogical purposes” (Gilquin & Granger 2010: 359). It is thus a very **hands-on approach** in using corpus linguistics in a pedagogical context or in education in general. The underlying concept of data-driven learning is providing “direct access to the data so that the learner can take part in building his or her own profiles of meanings and uses” (Johns 1991: 30). The idea is to make the learner a “linguistic researcher” (Johns 2002: 108), a “language traveller” (Bernardini 2001: 227) or a “language detective”, referring to Johns’ famous quote: “Every student a

Sherlock Holmes!” (Johns 1997: 101). DDL is thus an **inductive way of instruction** as the learner is in the center of the learning process and the learner formulates hypothesis and proves them valid or not and draws conclusions from their findings.

Advantages and limitations

One of the main advantages of DDL is that it “brings authenticity into the classroom” (Gilquin & Granger 2010: 359) by using corpora which depict authentic language. There is real language evidence of how language is used, especially how words or patterns are used. DDL is typically used in the context of teaching linguistic competence, including vocab and grammar patterns as it “contribute[s] to vocabulary expansion or heightened awareness of language patterns” (Gilquin & Granger 2010: 359). In terms of learning skills, O’Sullivan claims that DDL provides a range of different applications and analysis techniques such as “predicting, observing, noticing, thinking, reasoning, analysing, interpreting, reflecting, exploring, making inferences (inductively or deductively), focusing, guessing, comparing, differentiating, theorising, hypothesising and verifying” (O’Sullivan 2007: 277). There are also social learning benefits as DDL involves teacher-learner interaction in giving advice during the research process but also in learner-learner interaction when collaboratively researching and discussing the data with peers (O’Keeffe et al. 2007).

As it is in inductive approach and the **learner is a researcher**, DDL has an “element of discovery” (Gilquin & Granger 2010: 359). This approach leads to a higher motivation among the students and is simply more fun (Gilquin & Granger 2010: 359). There is some evidence that DDL is also effective in terms of learning output, even though more research needs to be conducted in order to gain more empirical evidence of the effectiveness of DDL (Mukherjee 2002; Gilquin & Granger 2010). So far, the effect relies on subjective observation and informal testing (Bernardini 2001: 247; Mukherjee 2006: 21). Many researchers see DDL as a very promising approach (Jablonkai & Csomay et al. 2023; Pérez-Paredes 2019; Römer 2011; Gilquin & Granger 2010; O’Keeffe et al. 2007; Mukherjee 2002; Hunston

2002). Cheng et al. (2003) for example found that in a university context, 80% of undergraduate students found corpus projects useful while Kennedy and Micelli (2001: 80) discovered “mixed feelings”. On the one hand, the students gained more confidence, but on the other hand, they found corpus work time-consuming and frustrating (Kennedy & Micelli 2001: 80).

There are in fact a few **limitations** which should be mentioned in this context. First of all, the teachers need to be educated very well and develop a profound “corpus literacy” (Mukherjee 2002), which means that they must have fundamental knowledge of corpus linguistics as in what kind of corpora exists and how to use corpus tools and techniques before integrating them in the classroom (Mukherjee 2002: 179). Another problem arises in what to present. The teachers could let the learners use a corpus by themselves, but then they could draw wrong conclusions from examples which they collected in a wrong way. If the material is pre-selected by the teachers, this reduces authenticity and motivation for the learners and it is also more time-consuming for the teachers. Also, using corpora in the classroom needs technical equipment which the school must either provide or be willing to buy. Not every school can afford this, which causes financial problems (Hadley 2002: 110).

However, these limitations could be overcome quite easily and are not too limiting if the impact of corpus linguistics was acknowledged more. If teachers are well-equipped with “corpus literacy” and know how to use corpus tools because they encountered corpus linguistics at university and in teacher training, less work needs to go into further educating teachers “on the job”. Letting students work collaboratively lowers the number of computers needed and also the cost for software and tools. It would be less time-consuming if more ready-made material was provided (see for example Römer 2011, Gilquin and Granger 2010).

There are for example new **open educational resources** (OERs) with **hands-on material** which were published in recent years (see Le Foll et al. 2021; and Pinto et al. 2023), which is a good start. Some of these aspects will be covered in greater detail in the following section in the context of teacher education. These limitations can be discussed in light of the “Mehrwerttheorie” (see section 2.1.6.)

and disregard the benefits and the learning effects data-driven learning and the integration of digital tools have in the EFL classroom. Data-driven learning and the use of corpora and corpus tools contributes to data literacy and establishes digital and future skills.

With these limitations in mind and being aware of them, I would argue that the use of DDL and corpora in language teaching is extremely useful. Gilquin and Granger also conclude that “on the basis of some of the claims found in literature, there are encouraging signs that DDL may deserve a place in the language classroom” (Gilquin & Granger 2010: 365). Most of the critique stems from a “Mehrwerttheorie”-viewpoint with criticizing the added value and the lack of proof for a better learning outcome. However, if DDL is seen with regard to promoting **digital skills** and working towards a culture of digitality, there is no doubt that it contributes to it. DDL is a very promising approach of authenticity and inductive teaching methods that puts the learner in the center of the learning process, which promotes effective language learning. Above all, data-driven learning with all advantages mentioned above, can contribute to data literacy and digital skills, which is a crucial element of the modern EFL classroom in the 21st century.

2.2.2.4. HOW CORPORA CAN BE USED IN THE CLASSROOM

There are several ways to integrate corpora into EFL learning and establish corpus-based activities. The following chart gives an overview on some possible applications (adapted from Mukherjee 2002: 160):

	Content	Function
Corpus as resource	What is a corpus?	What are advantages of corpus data?
Corpus methods	How can a corpus be analyzed? What are basic corpus tools and techniques?	What findings can be drawn from corpus analyses?
Using corpora	Corpus-based activities	

FIGURE 22: APPLICATIONS OF CORPORA IN THE CLASSROOM (ADAPTED FROM MUKHERJEE 2002: 160)

As the categories in figure 22 will be explored in greater detail in the following, I will use the categories outlining how corpora can be used in teaching introduced by Aston (2002) and Leech (1997) mentioned earlier – teaching about corpora, exploiting corpora to teach languages and teaching to exploit corpora – for structural purposes.

Teaching about corpora

The aim of this category is to give the learners a **basic understanding of corpora**. Teaching activities are centered around questions about what corpora are and what advantages corpora have. A basic definition of what a corpus is and information of how corpora are categorized, like written versus spoken corpora, should be the subject of teaching. Giving examples of how to use some corpora and providing a list with corpora is very useful for the learners. Some basic principles of how corpora are built and what criteria for designing corpora exist could also be the content of a lesson. These include for instance the concept of representativeness, size of corpora and balance. The fact that frequency plays an important role in corpus linguistics and when working with corpora should of course also be explained. As far as advantages of working with corpora are concerned, the fact that corpora provide empirical evidence of language in use and authenticity of the data are crucial. Additionally, it should be explored that corpora have a wide range of applications and that even basic dictionaries and grammars, which the learners probably consult, are based on corpus data. A conscious way of working with corpora should be encouraged and therefore, limitations of corpora should also be discussed. This is a very teacher-centered approach as learners usually do not use corpora themselves at this stage.

Exploiting corpora to teach languages

The aim of exploiting corpora to teach languages is to explain basic corpus methods to the learners so that they are able in the next step, to **exploit corpora on their own**. From a content perspective, it should be taught how a corpus can be analyzed and thus basic corpus tools and techniques should be introduced and trained. Corpus tools like *Wordsmith* or *AntConc* or especially developed corpus tools

for learners like *SketchEngine* and *SKELL* could be explored. Depending on how advanced the learners are, the teachers could in a first step demonstrate the functions of these tools and conduct some basic searches. In a second step, learners could use these tools on their own and replicate the searches the teachers conducted in order to get used to the functionality of the tools. Examples for corpus analyses could be exploiting word lists in order to recapitulate the frequency principle of corpus linguistics. Different word lists could be compared to each other and keywords could be identified. Learning how to “read” concordances and finding patterns would also be an important corpus technique which could be exercised in this category. Sinclair (2003) formulated seven steps how to analyze concordances, which could be helpful as guidance through the analysis process:

1. Initiate
2. Interpret
3. Consolidate
4. Report
5. Recycle
6. Result
7. Repeat

These steps could be used as a guide for learners how to analyze concordances. On a functional level, the learners should be able to draw conclusions from analyses which could be conducted. In order to practice how to use corpus tools, it is suggested that the teachers demonstrate and use the tools on a computer. As far as finding patterns and keywords is concerned, the teachers could either provide the learners with wordlists or the learners could exploit wordlists from corpora themselves. This depends on the level of proficiency and the methodology.

This category counts as data-driven learning as it “confronts the learner as directly as possible with the data” (Johns 2002: 108). The underlying principle in this category is that the students find patterns in a corpus and then make assumptions and generalizations about language. In doing so, they are using

corpus data in order to learn more about actual language in use and can integrate these findings into their own language use.

At this level, learners are more involved even though the tasks are guided by teachers. There is much teacher-learner interaction as teachers explain and demonstrate uses of tools and techniques, and learners can try it out themselves. However, a truly data-driven and learner-centered approach is only applied in the following category.

Teaching to exploit corpora

The aim of teaching to exploit corpora is that learners can **use corpora autonomously** and conduct corpus searches themselves. At this stage, they need to have fundamental knowledge of corpora and of basic corpus tools and techniques. It is also helpful if they are used to corpus tools and can use them autonomously. The learner turns into a researcher and has to conduct his own research.

Kennedy and Miceli came up with **four steps** for how to **investigate a corpus** (Kennedy & Miceli 2001: 82):

1. Formulate the question
2. Search strategy
3. Observe examples and select relevant examples
4. Draw conclusions

These are basic steps in research, but learners are usually not familiar with this kind of scientific research approach. Therefore, it is useful to provide them with this list or guide them through this process so that they know what to do and have a strategy on how to conduct the research.

In a first step, learners should come up with a **question**, ideally a question about language use. They could be inspired by earlier searches they conducted under teacher guidance. They could also come up with new questions based on problems they were facing during their learning experience. They

could for example investigate the difference in use between related words such as *can* and *may* or *do* and *make*. Altenberg and Granger (2001) suggest concordance exercises with high frequency verbs like *make* as they produce difficulties for the learners. Another typical search would be how prepositions are used with certain verbs, such as what prepositions follow *look* and what functions the different constructions have.

In a second step, the learners should be able to **select** the corpora they would like to investigate based on characteristics of the corpus. Looking for spoken language features in a written corpus for example is not very suitable. The second step also includes methodological decisions, for example which analyses should be carried out and why the respective kinds of analyses were chosen. Typically, quantitative and qualitative analyses should be combined. It is a good start to begin with frequency analysis such as word lists or KWIC analysis before carrying out collocation analyses or looking at concordances.

In the next step, learners should be able to **observe** examples and make decisions about which examples are relevant for the respective research question and which ones to ignore. Frequency typically draws the attention to keywords which should be investigated further. If a keyword had already been selected in the research question, for example if the difference between interesting and nice should be investigated, it would still be worth starting with a frequency analysis by comparing which word was used more frequently and in which context.

The last step is concerned with drawing **conclusions** from the examples which were presented and selected through the analysis process. The learners interpret the results by stating how specific words or patterns were used in the corpus. They are able to prove (or disprove) their hypothesis and thus formulate possible answers to their research questions. These conclusions drawn from language use in the corpus typically lead to generalizations about actual language use. It is worth also including statistical information so that generalizations solely based on quantitative analysis are possible, especially when working with smaller corpora. Typically, with an additional qualitative analysis,

statistical information is not necessarily needed and the qualitative analysis gives hints about strategies or decisions of actual language use.

Teaching to exploit corpora thus is a very learner-centered approach which clearly belongs into the category of data-driven learning. Not only communicative competence but also social skills and data literacy are important aspects of learning in this context. Learner-learner interaction makes up the most part of the lessons as the learners either work cooperatively together or they present their findings to their peers.

The positive impact of using corpora and corpus methods in the classroom on learners was claimed for example by Boulton (2009): “corpus samples led to more successful results than traditional pedagogical resources such as bilingual dictionaries or usage manuals” (Boulton 2009: 50). Granath echoed this positive impact by stating that “advanced students definitely benefit from working with corpora” (Granath 2009: 59). However, as mentioned above, more studies on the empirical effect of learner outcome need to be conducted, even though most of the teachers’ intuition and personal experiences would suggest to integrate CL in EFL teaching.

Corpus literacy in teacher education

As mentioned before, using corpora in the classroom starts earlier by raising awareness of data-driven learning and corpus linguistics among teachers in **teacher education**. Mukherjee, who conducted a survey among English language teachers in Germany, claims that “it is the teachers to whom particular attention should be paid in this process of popularization” (Mukherjee 2004: 244). His main findings are that 80% of the participants of his survey “had not come across corpus linguistics before” and that “virtually all (i.e. more than 95%) do think that English language teaching may profit in one way or another from the advent of corpora” (Mukherjee 2004: 242). The participants also seem to unconsciously use corpus linguistics already as “[p]aradoxically, most of the teachers who took part in the survey admitted using corpus-based resources, especially corpus-based dictionaries” (Mukherjee

2004: 243). According to his findings, teachers find data-driven learning and the use of corpora and corpus methods worth teaching and important. Thus, the majority of the teachers in this survey see the need and the advantages of using corpus linguistics and it should be integrated in teacher education at all stages – at university level, in teacher training in schools and in further education. Farr and Leńko-Szymańska (2023) propose a framework for the integration of corpus linguistics in teacher education consisting of three main elements:

- **Corpus literacy** (see Mukherjee 2006): What is a corpus? What can a corpus do? How can a corpus be analyzed? How can conclusions be drawn from this?
- **Corpus-based Language Pedagogy (CBLP)** as “the ability to integrate corpus linguistics technology into classroom language pedagogy” (Ma, Tang, & Lin 2021: 2734)
- **Corpus-based Reflective Practice**

They also draw the light to four studies and surveys among **English language teachers** (see Breyer 2009, Abdel-Latif 2020, Poole 2022 and Ma, Chiu, Lin & Mendoza 2023). These studies describe positive effects the teachers mentioned when using corpora in the classroom. The teachers were enthusiastic using corpora, they liked the way how new material evolved, they appreciate authentic data and they saw advantages of using corpora to teach vocabulary and grammar. However, they also mentioned negative aspects. Using corpora is very time-consuming due to a certain corpus-literacy that needs to be established, i.e. the teachers need to be educated as well as the learners. Technological equipment is also sometimes not available and data-driven learning works better with advanced learners so that it is not suited very well for beginner learners.

The teachers in these studies claimed that one of the biggest problems is the **lack of ready-made examples** which could easily be integrated in teaching. Ute Römer already suggested more ready-made material in her article published in 2011: “In the future, I would hope to see more publications [...] that contain ready-made exercises based on authentic speech and writing from different text types and language varieties and focused on language items that are of central importance and/or troublesome for learner” (Römer 2011: 216). The good news is that new ready-made material is being

developed and this is probably one of the biggest challenges for the next years and a deciding factor which either helps integrating CL in the classroom or not at all. Even though some OER for teachers with hands-on material and lesson plans, which are methodologically very useful, have been developed in recent years (see Le Foll et al. 2021 or Pinto et al. 2023), there is still much work that needs to be done. It is the teachers who need to be convinced – not only content-wise, but rather that corpora and corpus methods are easy to integrate – and also syllabus designers. We know from a didactic point of view that what is integrated in the syllabus or the curriculum also finds its way into teaching. Syllabus designers claim for digital literacy and in Germany they also suggest integrating digitization into the EFL classroom (see KMK 2016, KMK 2021), but they fail to draw the connection to data-driven learning and mentioning how valuable integrating corpora and corpus methods into teaching could be.

To sum up, corpus linguistics revolutionized the way language was analyzed and it has the potential to as well revolutionize EFL teaching. In this section, I looked at what corpora are and what makes their use in EFL teaching so attractive: Corpora display actual language in use and thus authentic language. In using corpus methods and corpus analyzing tools such as frequency analysis or concordancing, utterances about how frequently words or phrases are used and how words are used in context can be made. This kind of information can help learners in various ways for example in vocabulary acquisition – how similar words have different meanings and are used differently – or in terms of phraseology and grammar as to how a word is used in context and in combination with other words (see collocations). These empirical methods provide information about words and phrases and thus draw on the learners communicative competence. When it comes to teaching, there are three main ways of using corpora: teaching about corpora, exploiting corpora to teach and teaching to exploit corpora (Aston 2000). These ways differ in the target group – teachers and learners are addressed here – and in the learner or teacher involvement. Using corpora in the EFL classroom is considered data-driven learning and the learner turns into a researcher and is actively involved in the learning process. He or she conducts his or her own research by formulating research questions and conducting the

corpus analysis. Depending on how to integrate corpora in the classroom, technological equipment is needed and the teachers need to be educated on how to conduct corpus analysis. There are some limitations, but all in all, using data-driven learning and integrating corpora in the classroom has proven to have more advantages than limitations.

3. DATA AND METHODOLOGY

This chapter is concerned with the **dataset** – the corpus – that was used and especially compiled for this study as well as the **methodology** that was applied in the analysis and application of the corpus. Section 3.1. provides a detailed description of the design and the compilation of the corpus. Section 3.2. focuses on the methodological framework of this work and the mixed-methods and interdisciplinary approach that was applied in analyzing the corpus and applying the corpus in an EFL teaching context.

3.1. DATA

The dataset that builds the basis for this work is the **corpus TaCoCASE**, the Transatlantic Component of the CASE Project (Collet 2023). Since this corpus had especially been designed and compiled for this study (and my PhD project), the design of the corpus, the data collection and compilation process, along with the issues and problems involved, will be described in greater detail in the following sections. Some general numbers of the corpus will be presented and these numbers will be compared to ViMELF, the sister corpus of TaCoCASE. I will close this section by presenting some key benefits of TaCoCASE and thus emphasize its rich data set and its broad range of application.

3.1.1. TACoCASE, ViMELF AND THE CASE PROJECT

The **CASE project** has been conducted at Saarland University and Trier University of Applied Sciences since 2012 with the aim of compiling a corpus of spoken English. The English Linguistic Research Group at Saarland University is mainly in charge of the project: Prof. Dr. Stefan Diemer is the project coordinator and one of the principal investigators along with Dr. Marie-Louise Brunner, Dr. Selina Schmidt and myself. The main focus of the CASE project lies on investigating **video-mediated communication** in **intercultural encounters**, which means conversations between speakers with

different language and cultural backgrounds. Therefore, the research team decided to compile a corpus consisting of data from Skype conversations between speakers of English as a Lingua Franca (ELF) for ViMELF and later add a native component to the corpus (TaCoCASE).

Sub-corpora of the CASE project

More than 200 hours of Skype conversations have been collected so far in the CASE project. The two sub-corpora that have been derived from the project and have been published online are **ViMELF** and **TaCoCASE**. Both corpora can be accessed online¹⁴.

Due to a cooperation with Birmingham City University, ViMELF as well as TaCoCASE are integrated in WebCorp LSE¹⁵ and can thus be searched online (see also section 3.1.2.)

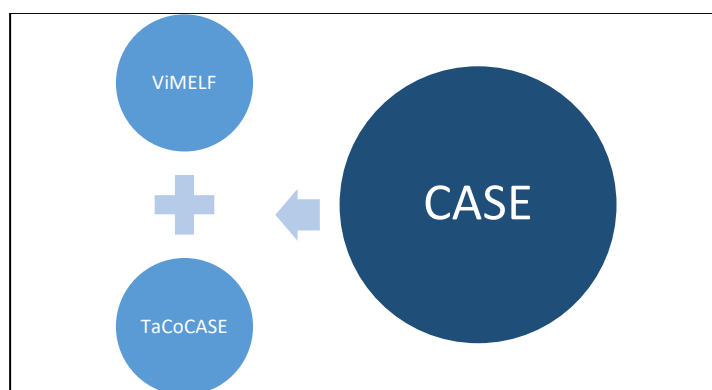


FIGURE 23: SUB-CORPORA OF THE CASE PROJECT: ViMELF AND TACOCASE

ViMELF (Diemer et al. 2018) is the corpus of video-mediated English as a Lingua Franca conversations and TaCoCASE (Collet 2023) is the corpus containing the transatlantic component of the CASE project.

¹⁴ <http://umwelt-campus.de/case/TaCoCASE>

¹⁵ <https://www.webcorp.org.uk/wcx/lse/corpora>

Depending on the research focus, researchers can either use ViMELF or TaCoCASE or combine the two corpora and thus have a larger dataset. The combination of the two is possible due to the fact that both corpora fulfill the general aim and guidelines of the CASE project and were compiled under similar conditions. I will not go into detail of the compilation of ViMELF or the CASE project here as the compilation process of TaCoCASE will be explained later in this chapter (see section 3.1.2.3.).

I will introduce the two sub-corpora of the CASE project, ViMELF and TaCoCASE, in the following and I will provide some basic information about these two corpora.

ViMELF

ViMELF (Diemer et al. 2018) was the first corpus derived from the CASE project that had officially been released in May 2018. The corpus, including video and audio material, as well as the transcripts, is available online and can be accessed by interested researchers upon request (<http://www.umwelt-campus.de/case>).

TABLE 8: ViMELF (DIEMER ET AL. 2018)

ViMELF (Video-mediated ELF conversations) (Diemer et al. 2018)	
Speakers	Non-native speakers of English
Conversations	20
Total length	744.5 min total, ca. 12.5 hours of conversations
Words/tokens	154,472
Participants	Total: 40 20 SB (German), 5 FL (Italian), 5 HE (Finnish), 5 ST (Spanish), 5 SF (Bulgarian)
Medium	Video/audio

ViMELF consists of 20 Skype conversations between students from different European universities: Saarland University (Germany), University of Bologna, Forlì (Italy), University of Santiago de Compostela (Spain), University of Helsinki (Finland) and St. Kliment Ohridski University, Sofia (Bulgaria). The participants were non-native speakers of English, which means that their first language is other than English (see section 2.1.3.1. and 3.1.2.2.3.). The total length of the conversations is about

12.5 hours and the number of words (tokens) is 154,472. Most of the recordings are video files, except a few audio recordings. In order to add a native speaker component to the CASE project and thus collect native speaker data as well, the CASE research team decided to compile another sub-corpus containing native speaker data under my coordination (as part of my PhD project): TaCoCASE.

The only main difference between ViMELF and TaCoCASE is the **speakers' language (and cultural) background**. ViMELF only consists of conversations between non-native speakers of English (NNS), who have different cultural backgrounds and use English as a lingua franca (ELF) as a means of communication (Seidlhofer 2011: 7). TaCoCASE, on the other hand, consists of conversations between non-native speakers as well as native speakers (NS) of English, who also have different cultural backgrounds and only to some extent different language backgrounds (NS-NNS conversations). I will not go further into detail as far as the ELF discussion is concerned and the difference between ELF, native and non-native speakers. The concept of native and non-native English speakers and the language model that should be used in an EFL context have already been discussed in sections 2.1.3.1. and 2.1.5.2.7. and the language background of the participants will be described in section 3.1.2.2.3. with regard to the speakers' (non-) nativeness. I will describe and explain the special setting of TaCoCASE in greater detail, which I will refer to as **ELF intercultural setting** (see section 3.1.2.2.4.).

3.1.2. TaCoCASE

TaCoCASE (Collet 2023), the transatlantic component of the CASE project, is another corpus of the CASE project. It was compiled at Saarland University and Trier University of Applied Sciences as part of my PhD project and was published online in September 2023. It consists of video-mediated conversations between English-speaking students from Britain, Germany and the United States, including visual and audio data. The conversations provide **authentic examples of intercultural encounters** between students with different language and cultural backgrounds, where communicative conventions and cultural concepts are negotiated. TaCoCASE can be accessed online

and it is open to all researchers who are interested in working with corpora. TaCoCASE can be applied and used in many different ways, which means that its existence contributes indirectly to research in various fields which go beyond the analysis and applications depicted in this work.

Access

TaCoCASE can be accessed for free and researched online via the following links:

- Access TaCoCASE: <http://umwelt-campus.de/case/TaCoCASE>
- Research TaCoCASE online via WebCorpLSE: <https://www.webcorp.org.uk/wcx/lse/corpora>

Due to a cooperation with Birmingham City University, TaCoCASE (and ViMELF) is integrated in the corpus search engine **WebCorp LSE**, which allows to conduct corpus searches online. Figure 24 and figure 25 are screenshots from the corpus search engine WebCorp LSE.

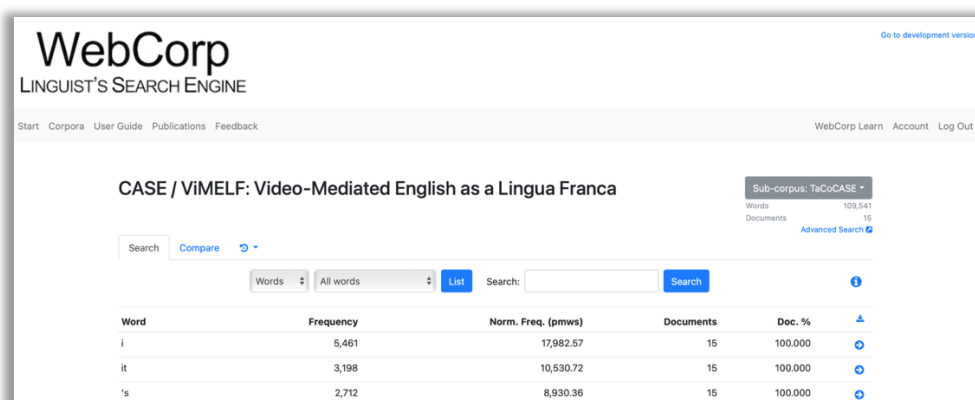


FIGURE 24: SCREENSHOT FROM WEBCORP (I)

FIGURE 25: SCREENSHOT FROM WEBCORP (II)

Using Webcorp LSE (WebCorp Linguist's Search Engine) means that most of the common corpus linguistics searches can be performed online, such as frequency, collocation or n-grams searches, wildcard searches and part-of-speech searches (see section 2.2.1.3.). The results can be displayed in concordance lines and in a KWIC format. It is also possible to display collocation tables and sort the data by statistical values or the criteria of interest. With the advanced search, comparisons between words or lemmas are possible, which is a very useful tool in the EFL classroom (see section 5.5.1.). WebCorp LSE could be used by the learners themselves in DDL activities (see section 2.2.2.3.; for an example see section 5.5.1.) as it functions like most corpus analyzing tools, such as AntConc and Wordsmith.

Besides using the web interface of WebCorp to analyze the corpus, it is also possible to use the data files of the corpus and work with them. The whole corpus could be loaded into AntConc or Wordsmith and analyzed with these corpus tools. The choice of tool depends on personal preferences and on the researcher's or teacher's scope or aim. Most of the corpora are only available as data files and are not integrated in a web-based search tool. Having access to TaCoCASE in both ways – WebCorp as well as the data files – is a benefit for users and makes the corpus easy to access and use.

3.1.2.1. CORPUS DESIGN

Collecting and displaying video-mediated spoken English language data, the general aim of the CASE project, is also the aim of TaCoCASE. As mentioned above, TaCoCASE contains the native component of the CASE project and thus serves as reference material to ViMELF, which is an ELF corpus or learner corpus. The main reason for compiling another corpus was to be able to separate native and non-native speakers and look at differences in language use between these speaker groups. This differentiation is traditionally made in the field of EFL teaching, in order to compare the language of native and non-native English speakers. With new concepts and new models of what kind of speaker models learners should aim for (see discussion in section 2.1.5.2.7.), this differentiation between native and non-native speakers is more and more questioned and it seems that the focus on native or non-nativeness seems not to be as central any more as it used to be – also taking the rising importance of ELF speakers into account (Byram 2021). In fact, the results in the analysis chapter (see chapter 4) suggest that, in TaCoCASE, the language between native and non-native speakers differs not too much and that other factors such as personal (lexical) preferences and having strategies and competences to communicate in intercultural situations seems to be more important. In order to see differences or similarities between native and non-native speakers though, native speaker data is still necessary and a separation between native and non-native speakers should be possible. Therefore, a native component to the CASE project was added with the data in TaCoCASE.

Aim and Rationale of TaCoCASE

My research focus is primarily on **EFL teaching** and English didactics (or EFL teaching methodology) and in particular how corpora can be used to improve English language teaching and learning. Corpora are not very present in language teaching (see Römer 2011; Mukherjee 2004; Jablonkai & Csomay et al. 2023) and in classrooms in general. Spoken corpora are rare in EFL teaching even though researchers usually point out that the applications of corpora in teaching and in the classroom are

manifold and teachers, students and curriculum designers (Römer 2011, Mukherjee 2004; O’Keeffe et al. 2007; Farr & Leńko-Szymańska 2023) would benefit. The benefits of using corpora in teaching have been discussed in section 2.2.2. in greater detail. The main advantage of using corpora in general, and TaCoCASE in particular, is that corpora represent authentic language use (Hunston 2002; O’Keeffe et al. 2007; McEnery & Hardie 2012). The language that the learners are confronted with is not always very authentic. There seems to be an “authenticity gap” between the language that is depicted in textbooks and language use in the real world (see section 2.2.2.1.). Corpora provide real-world authentic examples, which can be very motivational for learners (see Gilmore 2007; Hunston 2002; O’Keeffe et al. 2007).

In TaCoCASE and the whole CASE project, the language that is depicted is very close to the students’ own language or use of language (similar age and topics) and close to their **real world experience**. One main didactic principle and a factor that is prescribed in the CEFR as well as the English syllabi of the federal states in Germany is that the content or the material presented in the EFL classroom should be close to the students’ world of experience (see sections 2.1.1. and 2.1.2.). This leads to a higher motivation for learning (Deci & Ryan 1993) and in the case of TaCoCASE, teachers can use the corpus or excerpts from the conversations as teaching material directly in the classroom (see various applications and teaching tasks in chapter 5). Examples from TaCoCASE can serve as examples of actual language in use and thus brings more authenticity in the classroom.

TaCoCASE is a **rich source** for application in the EFL classroom and it can be used in particular to establish communicative as well as intercultural competences, as suggested by the CEFR and the education standards (see chapter 5). It can also be used to promote digital competences in EFL teaching with data-driven learning activities, which not only upgrades EFL teaching, but is also demanded by EU and German education standards and norms. Thus, using examples from TaCoCASE and using the corpus in EFL teaching promotes several competences at once and suggests a new, modern and data-driven way of language teaching.

Besides the fact that TaCoCASE is very well suited for teaching and corpus-based pedagogy (CBP), TaCoCASE is a very versatile corpus that allows research in many different research fields. Corpora can be used in many ways and in many fields to investigate language in use (see McEnery & Hardie 2012; Hunston 2002; see also section 2.2.1.). TaCoCASE, for example could also be used for investigating various **linguistic aspects** or linguistic elements such as individual words, patterns of language, lexical bundles, discourse markers, pragmatic markers or non-verbal elements like laughter or gestures. Depending on the research focus, the data provides many possibilities and serves as a rich source to investigate actual and authentic language in use. This broad range of applications was one of the main aims of compiling TaCoCASE and the diverse applications are only possible due to the special set-up of the corpus and the special transcription scheme (see section 3.1.2.3.2.). A list of benefits of TaCoCASE and its applications in the fields of corpus linguistics and linguistics in general are described at the end of this section (see section 3.1.3).

TaCoCASE is also designed for **application in the field of intercultural communication**, taking the different cultural backgrounds of the participants into account. It could be investigated in which cases intercultural communication is successful, what kind of strategies the speakers use to negotiate meaning and to explain cultural concepts. The results from this (inter)cultural analysis or selected excerpts from the data can be used to promote intercultural competence and cultural sensitivity. Looking at examples and making learners aware of cultural differences and similarities and what the participants do to overcome these differences and to what extent they value similarities contributes and expands the learners' own intercultural competence. This means, on the one hand, investigating culture-specific aspects which belong specifically to the German, British or U.S.-American culture, like for example traditional food or cultural traditions, and on the other hand, investigating aspects that are more of a global or trans-cultural nature, such as aspects of popular culture.

In sum, my aim was that to create a corpus of authentic CMC conversations which represent intercultural encounters between students with different language and cultural backgrounds. The corpus serves as a rich data source for researchers and is applicable in many different research areas.

Examples from the corpus can be adopted in EFL teaching and provide authentic language data that could be used in the EFL classroom. TaCoCASE is particularly suited to conduct research in linguistics and intercultural communication and the application of the corpus is beneficial in the field of EFL teaching.

Selected criteria for compilation

O’Keeffe and colleagues describe **five steps** to build a spoken corpus (O’Keeffe et al. 2007: 5 ff):

1. Create a design rationale – what should your corpus represent? How much data?
2. Record data
3. Transcribe recordings and save as text files
4. Database texts (,header’ information, gender, age, number etc.)
5. Check transcription

These steps represent a very common way to compile a spoken corpus and these steps were also applied in the compilation of TaCoCASE. The first step, the **design rationale**, has already been explained. Steps 2 to 5 will be explained in greater detail in section 3.1.2.3., which focuses on the compilation process of the corpus, in particular section 3.1.2.3.1. (recording phase), section 3.1.2.3.2. (transcription phase) and section 3.1.2.3.3. (anonymization and correction phase).

There are several factors and criteria for designing a corpus, such as size, balance and representativeness (see Hunston 2002; see also section 2.2.1.1.). TaCoCASE is a spoken corpus and is quite small compared to most written corpora, which is normal due to the very time-consuming compilation and transcription (see Diemer et al. 2016). It fits quite well to medium-sized spoken corpora with 10.5 hours of conversations and 140,003 words. TaCoCASE is balanced in terms of the number of subsets per speaker combinations (three subsets, five conversations each) and in terms of length of conversations and number of words (see section 3.2.2.2). TaCoCASE cannot be considered to be representative of ELF language or language of native or non-native speakers of English as the

number of speakers and the number of words is not large enough to allow generalizations. However, it is representative enough in order to make qualitative statements about language preferences or cultural aspects of the respective set of conversations by the respective set of speakers in this very context. Susan Hunston pointed out that “[i]t is a truism that a corpus is neither good nor bad in itself, but suited or not suited to a particular purpose” (Hunston 2002: 26). Following this quote, I argue that TaCoCASE is very well suited to the purpose and the aim of this dissertation while still being applicable to many other fields and research interests.

The results depicted in chapter 4, for example, provide serious evidence that the native and non-native speakers differ not too much in terms of language use and that native speakers also use communication strategies in intercultural encounters that were found and established in ELF research. This means that, on the one hand, established strategies from other fields, such as ELF research, can be found in TaCoCASE. On the other hand, the findings of the analysis of TaCoCASE (see chapter 4) question traditional language models in EFL teaching. It seems that the traditional distinction between native and non-native speakers that is usually made in EFL teaching (see Mukherjee 2002; Byram 2021), does not fit to authentic language use in TaCoCASE and that in intercultural encounters, different strategies and rules exist (see Seidlhofer 2011; Mauranen 2012; Brunner 2021; Antonello 2022). Therefore, I suggest new models and new ways of teaching, taking the findings of the analysis into account (see chapters 4 and 5).

3.1.2.2. CORPUS INFORMATION

In order to gain a deeper understanding of TaCoCASE, this section provides some **basic information** on the corpus itself. TaCoCASE is a spoken CMC corpus, which consists of intercultural encounters between students with different language and cultural backgrounds. The conversations were held via Skype and the participants were students from Germany, Great Britain and the United States, aged between 15 and 35, who voluntarily took part in the project. All conversations are first encounters and

took place in a dyadic and informal setting without surveillance from the research team. The participants received a topic prompt but were free to talk about any topic of their choice. The conversations should last about thirty minutes, but the participants could continue the conversations as long as they wanted as long as the conversation proceeded in a natural way. Due to the specific set-up of the corpus, CMC features as well as pragmatic elements and multimodal features are important components of the corpus, alongside the fact that the corpus displays spoken English and authentic language in use.

3.1.2.2.1. GENERAL NUMBERS

TaCoCASE (Collet 2023) consists of about **10.5 hours** of conversations between native and non-native speakers of English. The conversations were recorded between 2016 and 2018. The recordings as well as the transcripts are available online (see section 3.1.2.). The following chart provides some key facts and figures about TaCoCASE.

TABLE 9: TACOCASE (COLLET 2023)

TaCoCASE (Transatlantic Component of the CASE project) (Collet 2023)	
Speakers	Native and non-native speakers of English
Conversations	15
Total length	650 min total, ca. 10.5 hours of conversations
Words/tokens	140,003
Participants	Total: 26 10 BI (British), 8 BO (US), 8 SB (German)
Medium	Video /audio

The 26 participants were students from the CASE partner universities Birmingham City University in Great Britain (BI) and Boise State University in Boise, Idaho in the United States (BO) as well as from Saarland University in Saarbrücken, Germany (SB).

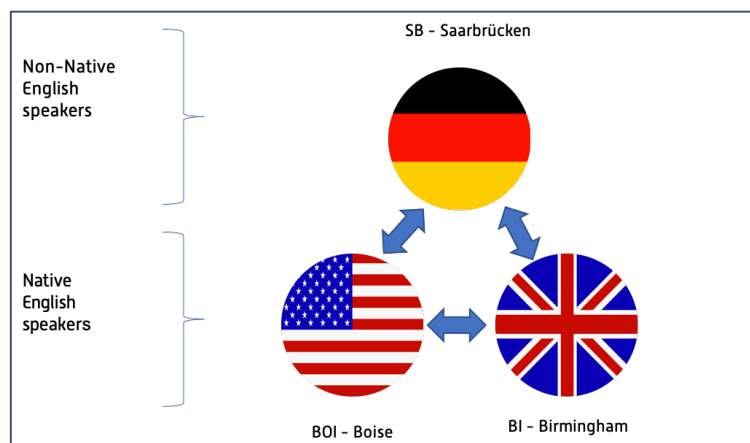


FIGURE 26: CONVERSATIONS – NS AND NNS IN TACOCASE

The corpus contains five conversations for each combination of students: BI-BO, SB-BI and SB-BO. The total length of the recordings is about 10.5 hours and the transcripts contain 140,003 words/tokens in general. The conversations were video files and only in two cases audio files.

Table 10 displays key facts about TaCoCASE and ViMELF in comparison to each other.

TABLE 10: COMPARISON OF VIMELF AND TACOCASE

	TaCoCASE (Transatlantic Component of the CASE project)	ViMELF (Video-mediated ELF conversations)
Speakers	Native and non-native speakers of English	Non-native speakers of English
Conversations	15	20
Total length	650 min total, ca. 10.5 hours of conversations	744.5 min total, ca. 12.5 hours of conversations
Words/tokens	140,003	154,472
Participants	Total: 26 10 BI (British), 8 BO (US), 8 SB (German)	Total: 40 20 SB (German), 5 FL (Italian), 5 HE (Finnish), 5 ST (Spanish), 5 SF (Bulgarian)
Medium	Video/audio	Video/audio

Based on these key facts, it can be stated that the two corpora are almost of the same size: ViMELF consists of 20 conversations, five more than TaCoCASE. The difference in total length of the conversations is two hours, while the number of words differs by 14,469 (about 10%). ViMELF has

more participants in total, but the conversation set-up is basically the same. Both corpora consist of five conversations per speaker combination: ViMELF consists of five conversations between students from SB and FL, SB-HE, SB-ST and SB-SF and TaCoCASE consists of five conversations between students from SB-BI, SB-BO and BI-BO. Both corpora contain mostly video conversations and only a few audio. In fact, TaCoCASE consists of 13 video conversations and two audio conversations.

3.1.2.2.2. CONVERSATIONS AND TOPICS

This section takes a closer look at the individual conversations of TaCoCASE. The participants held the conversations with their interlocutor via Skype. As the participants did not know each other and have not met before, the conversations can be considered **first encounters** between students from different cultural backgrounds. The participants were provided with the following **topic prompts** (see table 11):

TABLE 11: CONVERSATION TOPICS

1. "So, what are you studying?" – Course of studies and job prospects
2. "And what do you do all day?" – Life as a university student
3. "University here and there" – Different academic cultures
4. "The role and future of English" – Lingua Franca & global attitudes
5. "Should learning be virtual?" – Mobile learning, online courses: pros & cons
6. "How do you celebrate?" – Cultures and traditions
7. "Let's talk about food!" Eating habits and preferences – what do you like to eat?
8. "What's on?" Popular Culture – Talk about your favorite TV show

These topic prompts served as starting point in order to get the conversation going. The topics, including course of studies, languages and learning, life and leisure time activities as well as cultural topics such as traditions, food or popular culture, are close to the students' real life. The participants were supposed to start with the given topic, but they were free to choose what they would like to talk about during the conversation, which means that they could switch topics during the conversation. Most of the participants did in fact start off with their allocated topic and continued to talk about other

topics or related issues during the course of the conversation. Some participants even covered more than one topic from the list. In terms of conversation length, the participants were asked to hold the conversation for at least 30 minutes, but they could go on for as long as they wanted. We made it very clear that the conversation should be as natural as possible without any break offs.

Table 12 provides an overview of the conversations, including information about the **distribution of topics, conversation length** and **number of words**. The table is sorted by topic number. Each conversation has a unique conversation ID (displayed in the second column), which consists of the topic number and the individual speaker IDs (see column three and four). Column five indicates whether the conversation is a video or audio and the last two columns show the length of each conversation and the number of words per conversation in the annotated version, respectively.

TABLE 12: TACOCASE CONVERSATIONS

Topic	Conversation ID	ID 1	ID 2	Video/Audio	Length	Number of words
1	01SB151BO05	SB151	BO05	Video	00:36:03	7639
2	02SB152BO06	SB152	BO06	Video	00:33:07	8614
2	02BI33BO10	BI33	BO10	Video	00:33:27	4733
3	03SB156BI24	SB156	BI24	Video	00:38:43	6654
3	03SB149BO03	SB149	BO03	Video	00:38:31	7912
4	04BO16BI40	BO16	BI40	Video	00:43:35	9452
5	05SB124BI14	SB124	BI14	Video	01:03:46	13096
6	06SB157BO11	SB157	BO11	Video	00:48:03	6653
6	06BO18BI38	BO18	BI38	Video/Audio	00:33:00	9735
7	07SB149BI31	SB149	BI31	Video	00:34:35	8082
7	07SB150BI32	SB150	BI32	Video	00:41:51	8678
7	07SB156BO10	SB156	BO10	Video	00:55:28	10691
8	08SB147BI29	SB147	BI29	Video	00:40:19	10144
8	08BO15BI36	BO15	BI36	Audio	00:50:00	17333
8	08BO18BI41	BO18	BI41	Video	00:59:00	12557

As displayed in table 12, most of the conversations are video files. Most of the conversations are roughly thirty minutes long and only two conversations are close to an hour or even a little longer than an hour, which leads to an **average conversation length of 43 minutes**.

Figure 27 depicts the **number of words per speaker combination** with regard to total number of words in the corpus. The BI-BO conversations make up 37% of words in the corpus, SB-BO conversations 28% and SB-BI conversations 35%. Figure 28 displays the **conversation length per speaker combination**. BI-BO conversations lasted 218 minutes, SB-BI conversations 216 minutes and SB-BO conversations 210 minutes.

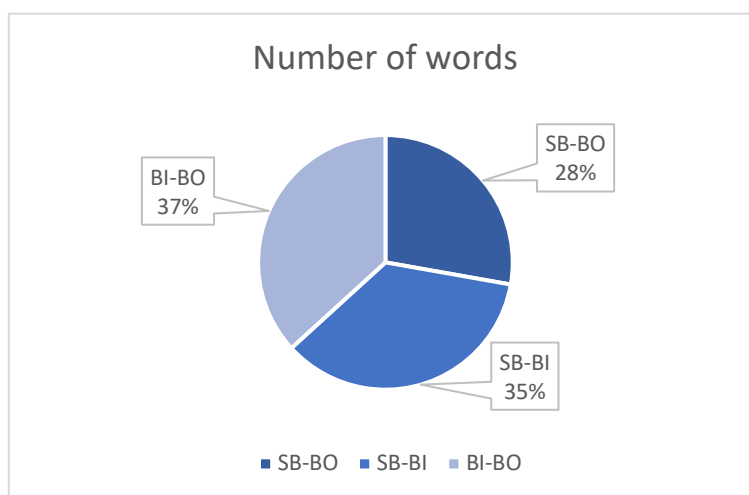


FIGURE 27: NUMBER OF WORDS PER SPEAKER COMBINATION

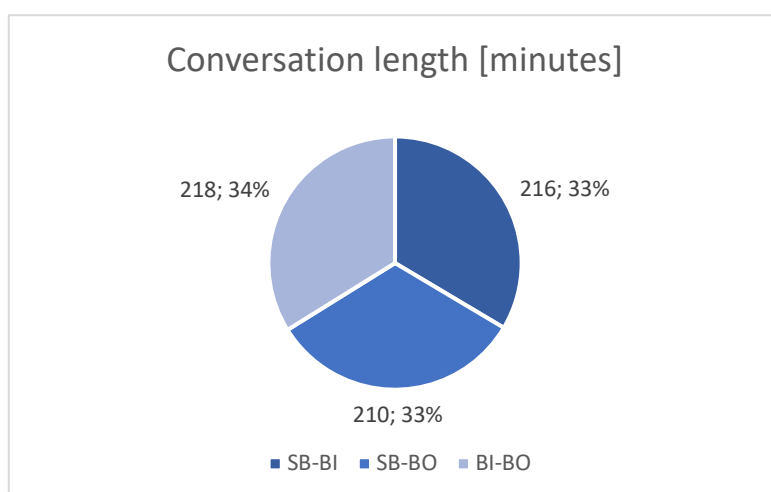


FIGURE 28: CONVERSATION LENGTH PER SPEAKER COMBINATION

The number of words and the average conversation length is quite equally distributed across speaker combinations, which is in fact by accident as there were no specifications about these factors. It can be stated that TaCoCASE is quite balanced in terms of speaker groups as far as the number of words and conversation length is concerned.

3.1.2.2.3. PARTICIPANTS

All 26 participants signed up for the CASE project with their instructor from our CASE partner university and took part in the recordings for TaCoCASE voluntarily. The participants were students and they were enrolled at the CASE partner universities at the time of the recording. Most of them were enrolled in language classes and particularly linguistics classes. Their age ranged from 15 to 35 years. The participants received information material and instructions about the project and the recording process (see section 3.1.2.3.1.) prior to the actual recording of the conversation.

Personal Information

26 participants took part in the TaCoCASE project and among them, eight participants identified as male and 18 as female¹⁶.

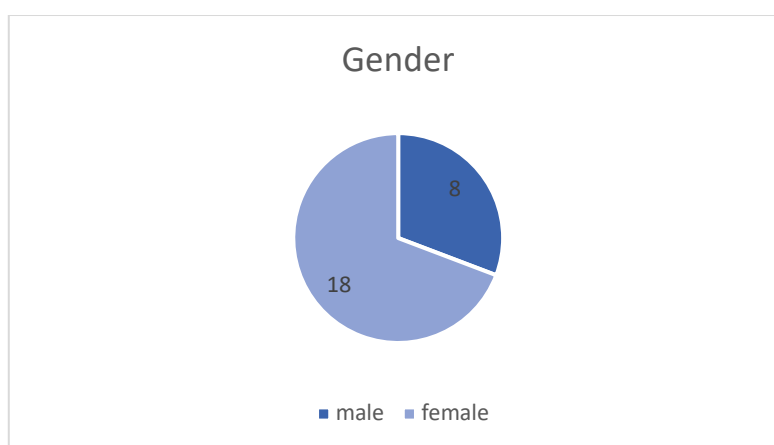


FIGURE 29: GENDER DISTRIBUTION IN TACOCASE

For the purpose of my studies and the corpus in general, it was not necessary to have a balanced number of female and male participants. As the participants were volunteers, we did not force a 50:50 distribution. It is also quite common for language classes, that there are more female students than

¹⁶ The questionnaire only provided the choice between male and female. A third option “diverse” was not offered. This is a limitation and would have to be adapted in future questionnaires in the project.

male students. This is a fact to keep in mind and be careful about when looking at gender-specific language use in the corpus.

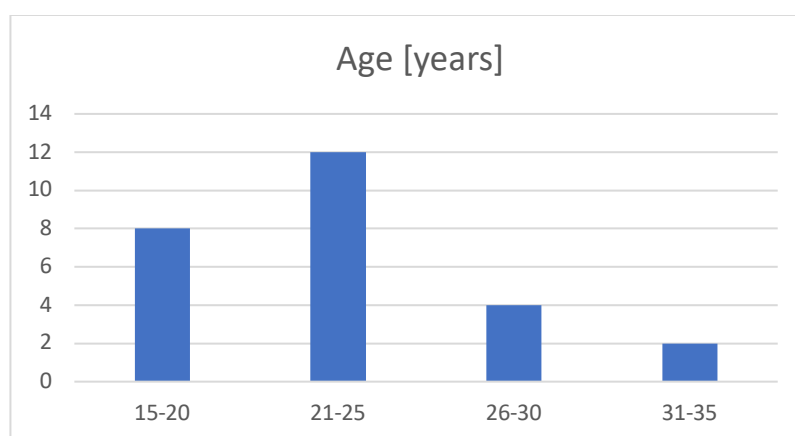


FIGURE 30 AGE DISTRIBUTION IN TACOCASE

Most of the participants (12 out of 26) were between 21 and 25 years old at the time of the recording. Eight participants were between 15 and 20 years old, four participants between 26 and 30 and only two participants were older than 30.

Each participant had a unique ID for reasons of anonymity. The ID contains information on the speakers affiliation: SB stands for Saarbrücken (Saarland University in Saarbrücken, Germany), BO for Boise (Boise State University in Boise, Idaho, U.S.) and BI for Birmingham (Birmingham City University, Birmingham, UK). The number which follows the letters represents the affiliation and the numbers stem from consecutive numbering of the participants from that one affiliation.

Table 13 provides an **overview of the personal information** mentioned above allocated to the individual speakers.

TABLE 13: PERSONAL INFORMATION OF THE PARTICIPANTS IN TACOCASE

ID	Gender	Age	Native Language
BI14	male	21-25	English
BI24	female	15-20	English
BI29	female	15-20	English
BI31	female	21-25	English
BI32	female	15-20	English
BI33	female	21-25	English
BI36	female	15-20	English
BI38	female	21-25	English
BI40	female	15-20	English
BI41	female	26-30	English
BO03	female	21-25	English
BO05	male	31-35	English
BO06	male	21-25	English
BO10	female	15-20	English
BO11	female	31-35	English
BO15	female	15-20	English
BO16	male	15-20	English
BO18	male	21-25	English
SB124	female	26-30	German
SB147	female	26-30	German
SB149	male	21-25	German
SB150	female	21-25	German
SB151	male	21-25	German
SB152	female	21-25	German
SB156	male	26-30	German
SB157	female	21-25	German

Language background

The participants from Boise (USA) and Birmingham (Great Britain) are **native English speakers** as they stated in the questionnaire that their native language is English. The participants from Saarbrücken (Germany) stated that their native language is German, which classifies them as **non-native speakers of English** (see section 2.1.3.1 for a detailed discussion of the concepts of native and non-native speakers). All participants could be classified as **intercultural speakers** in TaCoCASE (see Byram 2021; see also discussion of the concept in section 2.1.5.2.7) and the setting of TaCoCASE can be described as intercultural ELF setting, which is a setting characterized by mutual understanding and the creation of common ground (see section 3.1.2.2.4.).

The big question with the data in this work is whether the conversations in the corpus TaCoCASE with native speakers (NS- NNS) can be regarded as ELF speakers. Traditionally, only ESL and EFL speakers would count as ELF speakers. However, following Mauranen and Seidlhofer's definitions mentioned above, intercultural encounters of speakers of different language backgrounds who use English as a means of communication can be regarded as ELF conversations. Native speakers can be regarded as ELF speakers under certain circumstances, as they "do not share a first language with those who speak English as an additional language, they fit the definition [of an ELF speaker]. [...] Native/non-native dyadic interaction is the limiting case" (Mauranen 2012: 7). This means that all conversations in ViMELF count as ELF conversations and the ones where non-native speakers of English speak with native speakers of English in TaCoCASE, can also be regarded as ELF conversations. The native-native conversations in TaCoCASE between speakers from the UK and the US do not fall into this category of ELF conversations. Native speakers in TaCoCASE are speakers from the US and the UK and clearly belong to the inner circle in Kachru's model. All other speakers are learners of English from different countries in Europe, which makes them English as a foreign language speakers and members of the expanding circle.

As far as **language learning** is concerned, the answers from the questionnaires indicate that the non-native English speakers have been learning English for 12 years on average. Given that the participants are senior level university students and they completed the German school system, where English is usually taught from 5th or 6th grade until 12th or 13th grade, this number seems to be quite reasonable. All German participants spent time abroad in an English speaking country. This seems to be surprising at first, but it is indeed a requirement for English studies on master level at Saarland University. On average, the participants spent eight months abroad, while the shortest stay abroad was three months and one student spent 24 months abroad.

As far as the native English speakers are concerned, one third (6 out of 18) of the native speaker participants stated that they do not speak a foreign language. Twelve participants stated that they

know at least one foreign languages and in fact, half of them know exactly one foreign language, three participants know two foreign languages and three participants know three or more foreign languages.

3.1.2.2.4. SETTING

The setting of a corpus has an influence on the language that is depicted in the corpus is and it is very characteristic of the corpus. It is important to take a closer look at the setting of TaCoCASE in order to understand **special features** of the corpus and the **language use**. TaCoCASE is a spoken CMC corpus that consists of conversations between speakers with different language and cultural backgrounds. I argue that there are three main aspects that characterize the setting for TaCoCASE: Informal academic context, intercultural ELF environment and computer-mediated communication (CMC), which I will explain in the following paragraphs.

Informal academic spoken language

Spoken corpora and especially spoken corpora with a multimodal setup are still not very widespread in comparison to written corpora (see section 2.2.1.2.; see McEnery & Hardie 2012; Hunston 2002). Spoken language has many interesting features such as contractions or overlap as well as the use of pragmatic elements such as laughter. Video data displays non-verbal elements such as gestures, gaze and embodiment, which play an important role in communication and which contribute to multimodal meaning-making. In every-day communication, spoken language is used far more frequently than written language.

The language in TaCoCASE can be described as **informal academic spoken language**, following Laitinen's (2015) taxonomy. This means that it can be put in the same category as informal talks between supervisors and students (outside advisory sessions) and informal talk during academic breaks, talk at social academic settings or team talk. It is important to contrast it to established corpora

of academic spoken language, like for example MICASE (2002) or ELFA (2008), which usually consist of seminar discussions, recorded lectures or seminars in a more formal and official university setting. In TaCoCASE, the participants are students and they recorded the conversations in their free-time and with no supervision of the CASE team. They were free to choose when and where to record the conversations. Usually, the participants decided to record the conversation at home in a familiar environment. The register can be classified as informal, yet the conversations are set in an academic context. Academic in this case does not mean academic language per se but rather that the conversations were recorded in an academic environment. The participants are university students and they were recruited at university to take part in the project, usually in a course context. Some topic prompts are academic topics, regarding university itself or academic disciplines, study-related topics and academic cultures, ELF and CMC context, identity and (popular) culture. Other topics that the participants discussed include academic careers and research projects.

Intercultural ELF setting

The students in the CASE project are from Bulgaria, Italy, Spain, Finland, Great Britain and the United States. The students in TaCoCASE were recruited through the CASE project partners in the UK, the U.S. and our project group in Germany and they have different language backgrounds as well as different cultural backgrounds. This classifies their interaction as **intercultural encounter** or intercultural communication (see Schugk 2004 and Lüsebrink 2016; section 2.2.1.1.). The conversations can be considered ELF conversations as ELF is the “preferred option for cross-cultural communication” (see Seidlhofer 2003: 9). Native-native conversations for example cannot be regarded as ELF conversations, whereas NS-NNS conversations are the “limiting case” (Mauranen 2012: 7). In fact, I would argue that the setting is neither only intercultural communication nor only ELF communication per se. Only looking at intercultural aspects would disregard linguistic aspects in the NS-NNS conversations and only looking at ELF aspects would exclude the NS-NS conversations, which are still intercultural

communications as intercultural topics and aspects influence the conversation. Thus, it is a combination of the two, which I will refer to as **intercultural ELF setting**, describing a continuum between intercultural communication and ELF communication, including elements from both research areas, which results in a combined setting with combined characteristic aspects.

Intercultural ELF setting – as used in this work – means that English is used as a means of communication in a cross-cultural context. As Hua (2016: 173) puts it, the common goal in intercultural and ELF research is “exploring how participants with differing lingua-cultural backgrounds negotiate meaning in intercultural encounters (Hua 2016: 173) and this is also the goal in the TaCoCASE conversations. The conversations are characterized by “co-construction of meaning” (Mauranen et al. 2015: 404) and negotiation of meaning (see Hua 2015) in order to ensure mutual understanding as successful communication is the overall goal in these conversations. It is not only negotiation of meaning in a linguistic sense, but also in a cultural sense, meaning that cultural practices as well as linguistic forms are being negotiated (see Baker 2009, 2015). The speakers in the intercultural ELF conversations use a set of communicative strategies to achieve “mutual comprehension” (Mauranen 2012: 7) and reach the goal of successful communication (see Jenkins 2015a, Mauranen 2012). We see that it is necessary for the participants to establish common ground (see Dewey 2012: 45; also Pitzl 2010, Kaur 2009) as this cannot be taken for granted according to ELF research (Cogo & Dewey 2012: 44-46).

While the focus in intercultural communication research is usually on misunderstandings or critical incidents (see for example Lüsebrink 2016, Baker 2015), which arise through different cultural concepts and backgrounds, the perspective taken in the intercultural ELF conversations is more positive, looking at successful communication with a more co-operative and collaborative nature. The participants are on a “joint endeavour” (Kaur 2009: 40; see also Baker 2015) to create common ground, explain cultural concepts and negotiate meaning linguistically and culturally, which facilitates to mutual understanding. The conversations are furthermore characterized by “linguistic adaptability and creativity” (Jenkins 2015a: 45), which is usually found in ELF research, as the speakers do not care so

much about (language) mistakes and they are not seen as deficiencies, the key is mutual understanding without judgement of language proficiency. Thus, the intercultural ELF setting is a very co-operative setting, which contributes to linguistic and cultural understanding. This goes hand in hand with the argument that language and culture should be taught in combination (see Kramsch 1991, 1998a, Haß 2006; see also section 2.1.5.2.) and with regard to the proposed model of the intercultural speaker (see 2.1.5.2.7.)

CMC setting

Video-mediated communication plays an important role in people's lives and has gained an even bigger impact during the Covid-19 pandemic (see for example Koester 2022; Sindoni 2023). Since people use videocalls and video-mediated communication in a private and business context, we decided that it is worth collecting such kind of data and investigating it with TaCoCASE and the CASE project. People use video-mediated communication in many different situations, for example when catching up with family members or friends who live abroad or in work settings such as business meetings or study groups at university. The medium of choice is usually Skype or a software that works similar to it. We opted for Skype because most people are familiar with it and we assumed that it would be easy to handle for our participants and the recording software worked fine for the purposes of this project. This decision was made long before the pandemic and thus long before Zoom and Webex and other software programs or tools were in widespread use.

Video-mediated communication is considered to be part of **computer mediated communication (CMC)** like online chats and discussion forums (Herring 2001, Warschauer & Grimes 2007). CMC varies according to the technologies on which it is based and according to its contexts of use (Herring 2002: 111). Typically, there are two types of CMC: asynchronous and synchronous CMC. Asynchronous CMC means that the participants are not participating at the same time in the conversation, which would for example be the case for E-Mails. Usually, one E-Mail is sent out and the other participants receives it, reads it and then answers the E-mail whenever he or she finds the time. This back-and-forth could

take place very fast or it could take some time depending on the availability of people involved. Synchronous CMC means that the participants are present or available at the same time when they are communicating with each other. This counts for real-time chat as well as video-mediated communication like in the Skype conversations in TaCoCASE. The language use varies in synchronous and asynchronous communication, as in synchronous communication the participants have to spontaneously react and formulate an answer without having too much time to think about the formulation or possible expressions. In asynchronous communication, there is more time to think and then formulate an answer and this whole writing process of writing and correcting could take longer.

Relationships in CMC are close to the ones in face-to-face communication (Walther 1996). Apparently, the medium, in this case the computer, does not have an influence on the relationship between the interlocutors. This is very interesting in terms of interpersonal behavior of the participants in the corpus, which would, according to Walther's findings, be expected to be close to spoken face-to-face communication. Video and audio channels imitate features of face-to-face interactions, facilitating informal communication (Fish et al. 1993: 50f.) In using a computer as a medium to communicate with each other, people who are a long distance apart from each other still have the possibility to get in touch. CMC thus facilitates intercultural communication regardless of national boundaries (Gong 2005: 1). New ways of working and collaborating with each other are possible without travelling to the interlocutor or without having to meet the interlocutor face-to-face. However, "distance still matters" (Olson & Olson 2000: 173; see also Walther 2011 and Bradner & Mark 2002). The medium itself, the computer or other technological devices present difficulties if they do not work properly – due to limited internet access or other technical issues. These issues such as "delays, echoes, [and] incomprehensible audio quality" interfere successful interpersonal communication and interaction (Tang & Isaacs 1993: 193f.).

The conversations in TaCoCASE can be regarded as synchronous CMC conversations or video-mediated communications, which clearly show aspects of CMC communication, such as delays, echos, interruptions of recordings and many more. These elements are especially taken care of in the

transcription guidelines as CMC features were added in the guidelines for the CASE project (see 3.1.2.3.2.; see also Diemer et al. 2016).

The language in TaCoCASE is informal and since it is real-time conversation, spoken language features such as hesitation markers or overlap can be found in the conversations and they appear very frequently. Investigating language in a CMC setting is very interesting due to the mentioned features. Since CMC communication is becoming more popular and more common, it is important to be aware of chances but also limitations of CMC especially with regard to technical problems and being able to successfully communicate even if technical problems arise. Having a set of strategies and being able to cope with these issues is always very useful. The participants in TaCoCASE show different strategies how to cope with these issues (see sections 5.2. and 5.4.), which could be used in teaching and which can serve as examples in teaching material.

3.1.2.3. COMPILATION PROCESS

The compilation process of the corpus is mainly the **data collection process** of the conversations, which followed best practices and the established steps in compiling a spoken corpus (see Hunston 2002; O’Keeffe et al. 2007; Thompson 2005). This process of compiling TaCoCASE consists of three main phases: the **recording phase**, **transcription phase** and **anonymization and correction phase** (see figure 31). The phases will be described in more detail in the following sections. The process as a whole will be outlined briefly in this section.

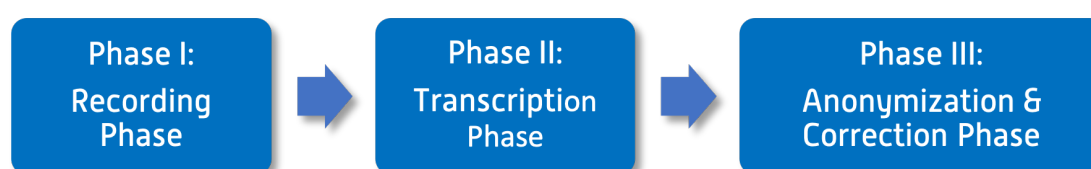


FIGURE 31: COMPILATION PHASES

The **recording phase** represents the beginning of the compilation process. The conversations were held and recorded during November 2016 and November 2018 by participants who volunteered to be part of the project. The participants recorded the conversations themselves and submitted the video files. The video and audio files and a metalinguistic questionnaire the participants had to fill in was then collected (see section 3.1.2.3.1.). The second phase is the **transcription phase**. In this phase, the recorded files were transcribed according to the CASE project transcription guidelines (see *CASE Transcription guidelines* 2017; see also section 3.1.2.3.2.). The third and last phase is the **anonymization and correction phase**. The video/audio files as well as the transcripts were anonymized according to the CASE project rules and standards, i.e. names and distinct locations were deleted or substituted (see section 3.1.2.3.3.). The transcripts passed the correction phase several times in order to guarantee a high quality of the final transcripts.

It is important to mention here, that the phases did not take place consecutively, which means completing one phase and then starting another phase. It was, for example, not possible to complete the recording phase for all recordings before starting to transcribe the data and thus entering the transcription phase, due to the fact that there were three recording phases. Therefore, it was easier to transcribe the data as it had been collected, which means that the individual recordings passed the phases as depicted in figure 2, but the process as a whole had several loops in between the phases.

3.1.2.3.1. RECORDING PHASE

The recording process can be divided into three phases: pre-recording, actual recording phase and post-recording. The processes in each phase will be described in the following.

Pre-Recording

Before the actual recording phase could take place, volunteers for the project had to be acquired. Since the CASE project has partners from the UK and the U.S., the contact to our volunteers was made through our colleagues from our CASE partner universities, Birmingham City University and Boise State University. Our colleagues advertised the project in their respective classes and students could sign up for the project. We then reached the students via E-Mail and we provided them with information on the project and the details about the recording. We provided them with and asked them to fill out the following files:

- Consent form
- Online questionnaire
- CASE Skype conversation guide
- Skype recorder instructions

The consent form provides legal proof that the participants give consent to be recorded and that the recordings can be used for research purposes (see Adolphs & Knight 2010: 42; Baude et al. 2010). As far as the online questionnaire is concerned, we used the **online questionnaire** that was used in the CASE project and that had been designed for non-native speakers for our non-native speakers and we designed a new online questionnaire for native speakers. While the focus in the non-native English speaker questionnaire is on English language learning and use of English, for example how long the participants have been learning English or to what extent they are exposed to or use English in their every-day life, the focus in the native English speaker questionnaire is on language acquisition and particularly on knowing other foreign languages.

The non-native questionnaires consist of three main parts: Part I consists of questions on personal data such as gender, nationality, age, native and other languages. Part II contains questions on English language learning, for example how long, how and where the participants have been learning English and part III is about English language use, for example where, how often and in which context the students speak English. While part I was the same for both speaker groups' questionnaires, part II (and III) differed. Part II in the native speaker questionnaire was concerned with the language background and time abroad and it contained questions like how many other languages the speakers know and if they spent time in a foreign language speaking country. Examples of both questionnaires can be found in the appendix (see appendix). The information in the questionnaire was anonymized, which meant that names and E-Mail address were deleted before publishing the background data online. For all other purposes, we only used the participant ID (see table 13).

Before the actual recordings were planned and took place, I provided the participants with instructions on how to conduct the recording in order to guarantee a high-quality recording. Two pieces of information were sent out to the participants: The **CASE Skype conversation guide** and the **Skype recorder instructions**. The CASE Skype conversation guide (see appendix) contains information on the requirements of the recording, such as length of conversation and set up of recording. It also entails a privacy notice, the list of topics and information on the submission process. Since the recordings in this

project took place under the same conditions as the recordings of the whole CASE project, the CASE conversation guide was used for the recordings in TaCoCASE as well. The Skype recorder instructions (see appendix) contained information on how to record the conversations, like for example, which recording software to use and what to keep in mind while setting up the recording (split-screen, volume and reduction of background noise).

After the volunteers gave consent to be part of the project, I compiled a list of participants and designed a conversation schedule. This means I paired the participants up randomly and assigned them a topic for their conversation. I contacted the participants individually and informed them about their conversation partner and their conversation topic. The participants then contacted each other individually, usually via E-Mail, and set up a date and time for the recording.

Recording

The actual recording took place at the responsibility of the participants. They set up their meetings individually and held the conversations at home or at university without any surveillance from our research team. The participants used their own computers, a recording software and usually the microphone and camera integrated in their computer, sometimes a remote camera or headphones and headsets. They called each other via Skype and held the conversation as previously arranged and following our guidelines. After the participants completed the conversation, they saved the video or audio file on their computer and sent the file to our research team.

The recording process started in November 2016 and took two years until November 2018, with three recording phases as illustrated in table 14.

TABLE 14: RECORDING PHASES

	Recording Phase I	Recording Phase II	Recording Phase III
Recording time	November 2016 - January 2017	November/ December 2017	November 2018
Total number of recordings	14	2	2
ID	01SB151BO05 02SB59BI27 02BI33BO10 02SB152BO06 03SB156BI24 03SB149BO03 04SB147BO01 05SB124BI14 06SB151BI33 06SB157BO11 07SB149BI31 07SB150BI32 07SB156BO10 08SB147BI29	08BO15BI36 06BO18BI38	04BO16BI40 08BO18BI41

It should be mentioned here that the majority of conversations had been recorded in recording phase I. The additional recording phases (recording phase II and III) were necessary because students dropped out of the project or because the data did not meet our requirements (see issues described in section 3.1.2.3.4.). This meant that the pre-recording and planning process had to be repeated again, which included finding new volunteers setting up and recording new conversations.

Post-Recording

As soon as I obtained the video or audio file, the files were stored in our CASE database. I did a quick cross-check whether the file was complete and the sound quality was okay right when I received the files. I also collected the online questionnaire and checked whether all necessary information was provided. I then stored the all files and the online questionnaires in our database.

3.1.2.3.2. TRANSCRIPTION PHASE

In the **transcription phase**, students from linguistics courses in our department at Saarland University and I transcribed the conversations. In some cases, the students who held the conversations were also the ones who transcribed it. It was important for us that we had language students and people who are familiar with the project so that they knew about the research goal and were aware of important aspects and features of the CASE project and its sub-corpora. Due to the several recording phases, there was not a single transcription phase, but there were three transcription phases. The transcription process took from January 2017 until April 2019 with the phases as depicted in table 15.

TABLE 15: TRANSCRIPTION PHASES

	Transcription Phase I	Transcription Phase II	Transcription Phase III
Transcription time	January 2017 - March 2017	December 2017 – March 2018	January 2019 – April 2019

All transcribers received instructions on how to create a transcript and they took part in a transcription workshop before they even received the anonymized video file. Additionally, we also provided them with transcription material: the transcription conventions and a two-page overview of the guidelines, which we found was helpful for quick checks of the abbreviations.

Transcription software

The **transcription software** that we suggested and that was mainly used was **F4/F5**. F4 runs on Windows computers and F5 is the Mac version. The software was already used for ViMELF and in the early phases of the CASE project and we found it very user-friendly which was also confirmed by our student transcribers.

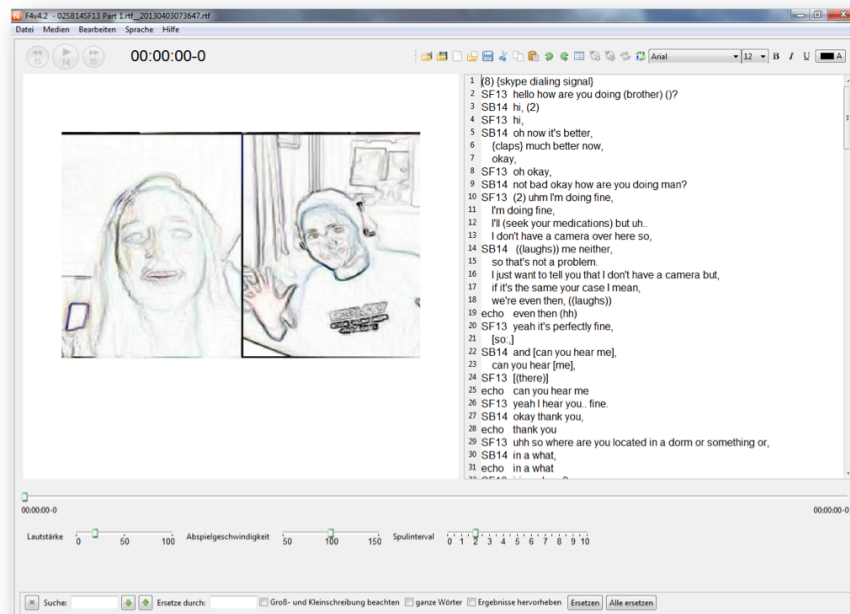


FIGURE 32: SCREENSHOT OF TRANSCRIPTION SOFTWARE

One of the main advantages of F4 and F5 is that the video is displayed on the one side of the screen and the transcript on the other (see figure 32). Especially when working with video material and gestures, the split screen of the software is very useful. The software also allows to create shortcuts, which makes transcription easier and faster as for example a shortcut for the speakers' ID or re-occurring non-verbal elements could be used.

Transcription conventions

Transcribing spoken language is always very difficult and a challenge for researchers and transcribers and as Thompson (2005) pointed out, spoken language corpora are "notoriously difficult to work with" (Thompson 2005: 59). In contrast to written language, spoken language has many features which make transcription very complex, such as contractions, repetition, overlap, or backchannelling. **The CASE transcription conventions** (see appendix) were designed taking the special language features of

spoken language in our data into account. The transcription of the TaCoCASE data followed the CASE transcription conventions, due to reasons of comparability and similarity.

Diemer et al. (2016) provide a very detailed description and discussion of challenges and issues in compiling and transcribing spoken corpora with special regard to the CASE project. I will briefly touch on the main issues and considerations for the CASE transcription guidelines, which were used for TaCoCASE, before taking a closer look at some examples of the transcription conventions.

In general, there are several possibilities to transcribe spoken language such as orthographic, pragmatic, prosodic, phonetic or phonemic transcription. Each of these ways of transcribing spoken language has their advantages and disadvantages. Orthographic transcription, for example, is most easily readable, but besides the pure orthographic description of what was said, there is not much information in the transcript. Since the video data in the CASE project is very rich, a pure orthographic transcription scheme seemed not to be sufficient. Prosodic transcription, as for example in the ICE corpora, is not very easily readable and is too specific in terms of prosody for our scope of studies (discourse, teaching).

Our solution therefore was to combine orthographic transcription with pragmatic elements as well as the key features of the data, CMC and ELF features (see Diemer et al 2016). Pragmatic elements include, following Levinson (1983), paralinguistic, non-verbal and prosodic cues (see table 17). Pragmatic elements, as well as CMC features and ELF features, which are very characteristic for the setting of TaCoCASE and the CASE project in general, were therefore added on top of the orthographic transcription. We used standard typeset characters to meet the requirement of readability and to make sure that transcripts are also readable with other operating systems and corpus analysis programs. Our general aim was to preserve as many contextual cues as possible so that the speech event is depicted as completely as possible while at the same time being easily readable and searchable with the help of corpus analysis tools.

The CASE transcription conventions are based on proposals and specific conventions by Dressler and Kreuz (2000) and Norrick (2013), which were both mainly developed for studying discourse. These proposals were further developed by the CASE team and adapted to the special features and setting of the corpus. While Dressler and Kreuz (2000) list for "paralinguistic" elements audible and visual cues, we opted to distinguish between audible cues, which are not lexical, like laughter, and non-verbal cues (gestures and gaze) (see Diemer et al. 2016). This means that laughter, a salient feature in our data, can be analyzed more in depth and especially with regard to their use, functions and role in conversation organization (see Jefferson et al. 1987; Spencer-Oatey 2002, Chafe 2007). New recommendations had to be added and were established due to technological advances and the special set-up of the data, including CMC features and ELF features (see Diemer et al. 2016). CMC features for example include overlap, echo and inference. ELF features include code-switching or non-standard language use.

In terms of visual arrangement of the data, we opted for a columnar arrangement, which makes the speaker turns visible (Thompson 2005, Edwards 1993). Each intonation unit is displayed in a separate line with the speaker ID at the beginning of the utterance to distinguish between speakers.

These considerations resulted in the *CASE transcription conventions*, which consist of four main categories: general conventions, specific conventions (prosody and pragmatics), ELF and CMC features as well as laughter. Some examples of the transcription conventions will be described in the following.

The **general conventions** include overlap, pauses, interruptions, liaisons and characters for unclear utterances (see table 16).

TABLE 16: EXAMPLES OF GENERAL CASE TRANSCRIPTION CONVENTIONS

Overlap	[]
Pauses	.. or ... or (x.0)
Interruptions, cutoff	–

Liaison	–
Unclear utterance	(utterance) or ()

Specific conventions were added taking prosodic and pragmatic features into account. Among these, we find characters for intonation and pitch, volume, and speed.

TABLE 17: EXAMPLES OF SPECIFIC CASE TRANSCRIPTION CONVENTIONS I

INTONATION & PITCH	
Rising	?
Falling	.
Continuing	,
Shift in voice	“ ”
VOLUME	
high	CAPS
low	° °
SPEED	
Prolonged sound	:
Quicker than surrounding discourse	> <
Slower than surrounding discourse	< >

Taking the special language into account, we also find ELF features in our data, which needed to be transcribed such as code-switching or non-standard pronunciation. **Pragmatic elements** like laughter or **CMC features** such as echo or interruption of recordings are also described in the transcription conventions (see table 18).

TABLE 18: EXAMPLES OF SPECIFIC CASE TRANSCRIPTION CONVENTIONS II

LINGUA FRANCA AND ELF FEATURES	
Code-switching	original transcription ((language (time in seconds)))
Non-standard pronunciation	corrected form ((/phonemic transcription/))
CMC FEATURES	
Echo	€€€ [€€€]
Interruption	{interruption of recording (time in seconds)}
LAUGHTER	
separate, loud laughter	((LAUGHS))
separate laughter	((laughs))
laughing while speaking	((laughing))
backchanneling laughter	((hehe))
chuckle, mouth closed	((chuckles))

We found in a study (Collet & Diemer 2014), which we conducted among 20 student transcribers, that it was hard for the transcribers to distinguish between the different kinds of laughter. As a result, we decided to provide audio examples of laughter so that the transcribers could listen in to these audio files and find out what kind of laughter represents the most accurate transcription.

The detailed **CASE transcription conventions** can be found online:

Transcription Conventions	https://www.umwelt-campus.de/fileadmin/Umwelt-Campus/SK-Weiterbildung/Dateien/CASE_Transcription_conventions_0817.pdf
Transcription Conventions (Overview)	https://www.umwelt-campus.de/fileadmin/Umwelt-Campus/SK-Weiterbildung/Dateien/Condensed_overview_CASE_transcription_conventions_0318.pdf
Taxonomy of non-verbal elements	https://www.umwelt-campus.de/fileadmin/Umwelt-Campus/SK-Weiterbildung/Dateien/Transcription_of_non-verbal_elements_in_CASE.pdf

3.1.2.3.3. ANONYMIZATION AND CORRECTION PHASE

For data protection and privacy reasons, the collected data needed to be **anonymized** so that it is not possible to retrace speakers' identity (see Adolphs & Knight 2010: 43; Baude et al. 2010: 66 ff). This includes the audio or video file itself but also the questionnaire and the transcript of the conversation. We followed established best practices as well as university guidelines in the CASE project (see for example Adolphs & Knight 2010; Baude et al. 2010; Wray et al. 1998; for a more detailed explanation see Diemer et al. 2016: 365ff). For the video file and the transcript, we decided to delete first names, last names and locations which are too close to one speaker or could give hints to their place of residence. Names were also deleted in the questionnaire and were substituted by the speakers' ID. It was very important to anonymize the video itself as it reveals the speakers' face and their background, which is very private if they recorded the conversations at home. We used a video filter from the current version of *Adobe Premiere Pro* so that only facial expressions or gestures are visible (see figures 33 and 34).



FIGURE 33: SCREENSHOT FROM CONVERSATION 06SB157BO11



FIGURE 34: SCREENSHOT FROM CONVERSATION 07SB149BI31

For TaCoCASE and ViMELF, video material is crucial as facial expressions and gestures are one key benefit of the data. Therefore, we decided to use a filter which still leaves contours but hides the speakers' identity. We guaranteed the participants anonymity and showed them beforehand how the video looked like with this special filter in order to be very transparent and trustworthy. Another reason for using this kind of filter – in contrast to most other corpora which leave the video unchanged – is that nowadays face recognition works very well and online tools that can remove filters are on the market, which could be used and reveal the identity of our speakers (Diemer et al. 2016: 365).

The **correction phase** included deleting errors in the transcript and the video data. The problem with the video was that the quality was sometimes not very good so that adjustments of volume or speed were needed. These changes were mainly done with the software *Adobe Premier Pro*. The correction of the transcripts was even more time-consuming because there were many levels of correction involved, such as fixing lexical errors, spelling errors, words or gestures which were not transcribed correctly. In addition, sometimes the technical form of the transcripts was not correct as for example a tab stop was not at the right place, which makes it hard for the corpus tools to read the transcripts. As the transcripts were made by different people, it was important to have one person, in this case myself, to do the final correction for reasons of similarity and consistency. This meant going back and forth from the video and the transcript and checking individual words as well as gestures for every single conversation. Despite the many issues and errors that had to be fixed, the whole correction process ensures high quality of the transcripts and the corpus in general, which make this phase a very essential element of the compilation process.

3.1.2.3.4. ISSUES AND CHALLENGES

As for the actual recording and the recording process as a whole, there are **four main issues** that I had to face. One major issue was a **high drop-out rate** of the participants. Some participants signed up for the project, but did not get in touch with their conversation partner and thus ended up not being part

of the project. This meant re-organizing the conversation pairs several times and repeating the planning process over and over again. Even some of the participants who wanted to be part of the project dropped out after they had a bad experience with one of their conversation partners. Thus, I usually started off with a higher number of students at the beginning, which decreased during the course of the project. This had an impact on the whole pre-recording process and the recording process took longer than originally planned.

Some participants had **problems with recording the conversation**, which led to massive support via E-Mail from the CASE team and in particular myself. Some students, however, did not ask for help and recorded the conversation in a wrong way without checking the guidelines and the material we provided them with. This led to conversations, which were not recorded side-by-side, but with a big window displaying one speaker and a small window with the other speaker instead. For the success of the project, the whole CASE team, however, decided to include these videos into our database, because gestures and facial expressions were still visible, which was important to us.

Some students had problems **using the transcription software**. Based on the experience with the first recording period, I decided to offer a transcription workshop prior to the second and third transcription phase. In this workshop, I presented the transcription guidelines, explained how to use the transcription software and did a sample transcription during the session with the students. This, in fact, led to better transcription quality in the second and third transcription phase.

The transcripts varied very much in terms of **quality** even though the guidelines were very clear. Some students missed important aspects, for example one student did not transcribe gestures at all or another student inserted two tab stops when only one tab stop should have been used. In the majority of the cases, the transcript was not wrong content-wise, but it was not correct from a technical angle or was not complete. It was very hard and very time-consuming to correct these transcripts. Despite these issues and problems, I was very aware that the participants were volunteers and handled them with great care in order to keep them in the project.

As far as **transcription** is concerned, the guidelines were very helpful during the process as they were especially adapted to the data in TaCoCASE. However, there are some **issues and limitations** that should be mentioned here. Despite many correction phases, the transcripts still vary to some extent in terms of accuracy, which means that there could still be some minor spelling errors or other mistakes that were not detected during the correction phases. The transcription scheme itself and the way the data is transcribed (see section 3.1.2.3.2.) is very suitable and provides enough information for most analyses. Phonetic transcription, for example, is not provided, which means that analysis in this field would not be possible with the transcripts. Since the video is also published, researchers who are interested in phonetic elements for example could consult the video files and then use their transcription scheme to conduct the analysis. Another aspect to keep in mind is that the transcripts are also not transcribed purely orthographic, which means that they are not too easy to read and researchers have to familiarize themselves with the transcription guidelines before using the data. Another problem that I discovered during the transcription process was consistency. The transcripts were made by different transcribers and thus every transcriber used a different standard – even though they were provided with the same guidelines. This counts for the variety of English, which was classified as British English in the standards, but sometimes transcribers used other standards out of habit. Non-standard forms like “wanna” were occasionally transcribed as the transcribers heard it instead of using the correct spelling “want to” (even though in spoken language this non-standard form would be appropriate). Length of intonation and arrangement of turns also varied across transcribers as well as the quantity of paralinguistic utterances.

These challenges were made up for in the correction phase. Since I corrected all the transcripts, a higher degree of accuracy and consistency was reached and the quality of the data is still very good. Decisions like not using phonetic transcription were made based on the underlying conditions (manual transcription, not too many transcribers), which are veritable decisions and very common in the compilation process. For the purpose of this work and the majority of research analyses, the transcription guidelines are very suitable and usable. Whenever more information is needed, the video

files can always be consulted. Even though the challenges were mostly made up for in the correction phase, it is still important to mention them to raise awareness and also provide suggestions so that researchers can learn how to deal with challenges during the compilation process.

3.1.3. BENEFITS OF TACOCASE

The previous sections of this chapter looked at the compilation process of TaCoCASE, which lead to the existence of a new corpus and a rich data source with some interesting features that allow a wide range of applications. This section looks at the **benefits of TaCoCASE**, which are displayed in the table below.

TABLE 19: BENEFITS OF TACOCASE

Key element	Description	Benefits for research / applications
Spoken language corpus	Spoken language features (e.g. overlap, intonation, hedges)	Authentic spoken language in use
CMC communication	CMC features (e.g. interruptions, echo), technical issues	Awareness of CMC features, strategies how to deal with interruptions or technical issues, building relationships in CMC communication (compared to face-to-face communication)
Multimodal set-up	Multimodal input: Audio and video: intonation, non-verbal elements (e.g. gestures, gaze, embodiment)	Different modes (text, video, gestures, environment); Multimodal meaning making and language use can be analyzed
Transcription	Different layers: orthographic, XML, etc. special guidelines: pragmatic elements (e.g. laughter), CMC elements (e.g. echo), ELF elements (e.g. code-switches), non-verbal elements (e.g. gestures)	Transcripts contain lots of (multimodal) information; depending on the research focus, different aspects can be focused on or layers can be disregarded
Authentic language use	Video recordings of conversations; no lab-setting	Actual language in use; authentic patterns, words and phrases
Informal register	Casual conversation via Skype between students; informal setting; prepared topic prompts but any topics can be discussed;	
Different speakers (native and non-native)	Native speakers and non-native speakers could be separated; TaCoCASE as reference corpus for	Different language use; speaker preferences; characteristics of NS and NNS (or ELF) language in use; comparisons between NS and NNS are possible

	ViMELF; ELF conversations (ViMELF plus TaCoCASE SB-BI and SB-BO)	
Intercultural encounters / intercultural communication	First encounters between students with different language and cultural backgrounds	Cultural aspects (culture-specific, transcultural), cultural awareness and sensitivity, promoting intercultural competence, strategies how to communicate in intercultural settings, strategies how to negotiate cultural concepts and meanings
TaCoCASE + ViMELF	Combination of the two corpora -> larger database	Comparisons across the two corpora (e.g. language, discourse marker use)
Access	Video and transcripts available online (for free); integration in WebCorplSE (searching the corpus online or comparing it with other corpora)	Corpus can be searched with corpus tools or online

This list of benefits of the corpus itself and its content offer many possibilities for researchers to use TaCoCASE in different research areas. Analyses and applications in the field of linguistics, applied linguistics as well as intercultural communication are possible. TaCoCASE can be used to investigate language in use and draw conclusions in different fields of research. It can be used to investigate how young people talk about every day topics and what kind of issues and topics they discuss. Excerpts of the conversations can be used as teaching or training material in various teaching contexts either at school, in higher education or in the business context either for improving communicative competence – when focusing on language features – or intercultural competence – when looking at how meanings and concepts are negotiated, for example in intercultural trainings.

In sum, TaCoCASE is a new corpus that had been carefully designed and compiled according to corpus linguistics standards and with special regard to the research aim in this work. It provides authentic CMC conversations between speakers with different linguacultural backgrounds and due to its set-up, it can be used and researched in a wide range of fields. I will explain in the following chapters (chapter 4 and 5) how the corpus can be used in a teaching context to promote communicative and intercultural competence and which linguistic and (inter)cultural aspects can be derived from the corpus. I will first describe the methodology that was used in order to conduct the analysis and the application in the next section.

3.2. METHODOLOGY

This chapter is concerned with the **methodology** applied in this study and in particular with the **methodological framework** used in the analysis and the application of the corpus data. In order to understand the methodological decisions, it is important to recall the overall aim of this work, which follows a didactic question: how corpora can be used to upgrade EFL teaching and to what extent corpora – and in particular TaCoCASE – can be used to promote communicative and intercultural competence.

Drawing on the main findings in DDL research (see section 2.2.2.), there is a gap between what corpus linguists and applied linguists suggest in research, namely integrating corpora and DDL activities in the classroom, and what actually happens in teaching, which is that corpora are not used very much. Chambers (2019) referred to this phenomenon as “research-practice gap” and Jablonkai and Csomay (2023: 3) tried to close this gap by collecting “the most recent findings in terms of what corpus data and corpus consultation can offer to the teaching practice” in the *The Routledge Handbook of Corpora and English Language teaching and learning* (Jablonkai & Csomay 2023). Many applied corpus linguists agree that corpora have great potential to revolutionize teaching and suggest to use corpora for pedagogical purposes (see for example Hunston 2002; Mukherjee 2002; O’Keeffe et al. 2007; Guilquin & Granger 2010, Römer 2011; see also Jablonkai & Csomay et al. 2023). This means corpora should be integrated directly in teaching, in the classroom itself, but also indirectly at other stages such as in teacher education or in material or textbook design (see Farr & Leńko-Szymańska 2023; Römer 2011, Mukherjee 2002; see also section 2.2.2.4.).

Despite the number of benefits that corpus-based pedagogy has and the suggestions made by applied corpus linguists, corpora seem not to make it into the classroom or at least their use in the classroom is not as common as it is suggested. Pérez-Paredes (2019) sees mainly two reasons for this problem: “Syllabus integration and a lack of contribution from language teachers other than researchers emerged as threats to the normalisation of corpora use” (Pérez-Paredes 2019: 36).

I agree with Pérez-Paredes and I see only two ways of integrating corpora in the classroom: either prescribing it in the syllabus or hoping that teachers would use it. The latter puts the responsibility on the teachers and despite positive attitudes towards using corpora in the classroom (see Poole 2020), most teachers are hesitant to use corpora for a number of reasons (DDL activities take time, lack of ready-made material etc.; see limitations discussed in section 2.2.2.4.). I argue that it is the policy makers' duty to integrate the use of corpora and DDL in the syllabus so that it will be used more commonly in the classroom as there is evidence in research that corpora have many benefits and could be very well used in teaching (see section 2.2.2.). If using corpora was prescribed in the syllabus, it would have a snowball effect in the sense that teachers would have to use it and would use it in the classroom more commonly. Additionally, corpora would be integrated in teacher training because teachers would have to know how to use corpora and material designers and textbook authors would integrate DDL activities in textbooks, which teachers still very much use to structure their lessons. The effect of integrating the use of corpora in the syllabus would be that the use of corpora as well as DDL activities would be used more often in the classroom, whereof EFL teaching in general would benefit. With this work, I intend to show that using corpora already fits to existing frameworks and standards, such as the CEFR, and that using corpora in general and TaCoCASE in particular promotes competences that are demanded by the CEFR. The competences are not new; using corpora – in this case TaCoCASE – and examples from the corpus to establish them is just a new approach that I suggest.

Having this overall aim in mind, suggesting new material and new methods of teaching, this work was designed accordingly, showing how TaCoCASE can be used to establish communicative and intercultural competence. I take the CEFR and the competences, objectives and operators that are described in the CEFR as a basis to establish communicative and intercultural competence. I established **research questions** for analyzing the data with regard to linguistics and (inter)cultural aspects and I will demonstrate how examples from the corpus fit to the competences mentioned in the CEFR – what I will refer to as **syllabus fit**.

Corpora have been frequently used as a basis for research and for pedagogical suggestions in the field of corpus-based pedagogy. What is new in this work is the approach to take the CEFR as a basis and to build the analysis and the application upon what is suggested in the CEFR in terms of communicative and intercultural competence. Therefore, I could not follow a well-established methodology for the overall aim. For the individual parts of the analysis – linguistic and (inter)cultural analysis – I picked well-established methods and approaches for the individual research questions and combined them in a mixed-methods approach on my own research having the CEFR and the competences as well as the overall aim in mind.

Methodological framework: a mixed-methods approach

As far as the methodological framework is concerned, I combined quantitative analyses and qualitative analysis in order to reach more fruitful results (Dörnyei 2007). I will explain the mixed-methods approach that I used in the analysis chapter from a more general perspective before going into more detail in this sub-chapter as well as at the beginning of the linguistic analysis section (see section 4.1.) and the (inter)cultural analysis section (see section 4.2.).

The **quantitative analyses** are based on common **corpus linguistics analyses**, such as frequency analysis, n-gram analysis or collocation analysis (see O’Keeffe et al. 2007; see also section 2.2.1.) in order to quantify the results and look at frequently used words and phrases and thus get an impression of the language used in the corpus. This quantitative analysis helps finding frequent and therefore relevant features of language and with corpus methods, these frequently used words can be easily retrieved from the corpus. Teubert (2005:1), for example emphasizes corpus linguistics’ very common use as a starting point in quantitative analyses by stating that using corpora is usually the “default resource” for linguists. Biber et al. ([1998] 2004: ix) go even further and point out the importance of using corpus linguistics as method as it “shed[s] new light on previously intractable research questions in linguistics” and thus emphasize its revolutionary character (see also Hunston 2002).

In the **qualitative analyses**, I combine mainly **discourse analysis** and **conversation analysis** along with specific sub-strands, such as multimodal discourse analysis, corpus-based discourse analysis or pragmatics. These qualitative methods are used in order to look at the examples more closely and find recurring patterns, unveiling underlying discourse structures and strategies as well as establishing models in a bottom-up process.

Discourse analysis (DA) is a very versatile method and can be applied in many different areas, which makes it also very suitable for my interdisciplinary work. Gee (2011) differentiated between discourse and Discourse (with a capital D): discourse (with a lower-case d) meaning “any instance of language-in-use or any stretch of spoken or written language” (Gee 2011: 205) and Discourse (with a capital D), which also includes “actions, interactions, ways of thinking, believing etc.” (Gee 2011: 201). Discourse analysis (DA) is thus more than looking at instances of language use in a text, but also including context and meaning in the analysis. In applied linguistics, for example, discourse analysis is usually used to identify “how people achieve certain communicative goals” or perform “communicative acts” (Paltridge 2012: 7), which is also an important aspect of my analysis in terms of communicative strategies. In terms of (inter)cultural or mediation strategies as well as explaining cultural concepts, discourse analysis is very suitable as it looks at “language and the social and cultural contexts” (McCarthy et al. 2013: 53). Brown and Yule (1983: 2) state that discourse analysis and pragmatics are very closely related. Pragmatics focuses on meaning and puts the speaker and listener in the center of the analysis by asking “[w]hat do they *mean* by that?” (Birner 2013: 1, emphasis in original). This is a very common question that will be used in my analysis as well in order to find out more about the intended meaning and underlying structures of discourse.

Multimodal discourse analysis focuses on different modes of communication (Scollon & Le Vine 2004: 1) and “language in combination with other resources” like gestures or sound (O’Halloran 2011: 120). This method can be used very well with TaCoCASE due to its multimodal set-up and it is very important to take multimodal features such as gestures into account when analyzing language. Non-verbal

elements in combination with verbal elements contribute to meaning-making, which is an important aspect of intercultural communication and facilitates mutual understanding (see Brunner 2021).

Corpus-based discourse analysis “look[s] at occurrence and re-occurrence of particular linguistic features to see how and where they occur in the discourse” (Paltridge 2012: 144) and thus simply combines corpus linguistics methods, such as frequency analysis, with discourse analysis. This combination of quantitative and qualitative methods sheds light on relevant features through frequency while at the same time looking at the data in a more close and qualitative way.

Looking even more closely at examples from a qualitative perspective means conducting a **conversation analysis (CA)**. The aim in conversation analysis is to “discover how participants understand and respond to one another in their turns at talk” (Hutchby and Wooffitt 2008: 12). Schegloff (2007: xiv), who is one of the main researchers who coined conversation analysis in the early stages besides Sacks and Jefferson, listed elements worth looking at in conversation analysis, for example turn-taking, sequence organization, trouble solving and overall structural organization (Schegloff 2007: xiv). These elements will be of interest when looking at how the participants explained cultural concepts and how they negotiated meaning. I established a model of a discourse structure for explaining cultural concepts in a bottom-up process, which means that I looked at individual examples and identified recurring elements in the discourse and came up with a general discourse structure. I also looked at turn-taking and sequence organization when looking at how the participants explained cultural concepts and how they perceived and discussed similarities and differences of cultural concepts. Thus, the overall guiding question in my conversation analysis can be summarized in one question: “Why that, in what way, right now?” (Seedhouse 2005: 167).

Taking a step back and looking at communicative, intercultural and digital competences, this already suggests an interdisciplinary approach as there are three fields that are crossed and combined: applied and corpus linguistics, intercultural communication and EFL teaching and didactics in general. The starting point is the analysis chapter, where applied and corpus linguistics methods are used in order

to analyze the language in the corpus (see linguistic analysis in section 4.1.) and how people act in intercultural encounters with regard to intercultural communication, negotiating meaning and explaining cultural concepts (see (inter)cultural analysis in section 4.2.), before the application of the corpus in the classroom and the syllabus fit in terms of communicative and intercultural competence is demonstrated (see application sections 5.1.-5.4.).

Aims of the analysis and application chapters

The aim of the analysis chapter (chapter 4) is to derive **information about the language used** in the corpus and use these findings as an **informed and evidence-based basis** for the application chapter (chapter 5). The analysis chapter provides information about linguistic and (inter)cultural aspects that appear in the data whereas the application chapter provides examples from the corpus that help promoting communicative and (inter)cultural competence in EFL teaching – based on the findings in the analysis chapter – and demonstrate the syllabus fit of TaCoCASE with regard to the competences mentioned in the CEFR. It is to mention here that the analysis and application chapters are logically interconnected as only aspects are analyzed which are of interest for the application, which means linguistic and pragmatic aspects as well as intercultural topics and discourse as mentioned before. These aspects are taken from the CEFR and I chose them because the CEFR lists them as important aspects of communicative and intercultural competence.

In the application chapter, **examples** from the corpus are presented that can be associated with the competences of interest – communicative and intercultural competence. These examples are presented in order to demonstrate the syllabus fit and they could be used in teaching as they match the required objectives and criteria in the CEFR. An even further and more concrete application is illustrated with case studies. The case studies present **hands-on material** and suggest lesson plans on how data-driven learning and the use of corpora can help teaching aspects of communicative and intercultural competence in the EFL classroom.

It is clear that due to the overall didactic aim, the linguistic and intercultural analysis and the EFL teaching approach in the application part, this work follows an interdisciplinary approach. The borders between the field of linguistics, applied and corpus linguistics, EFL teaching, cultural studies and intercultural communication are continuously crossed and sometimes methods from two fields are used in combination in order to reach more fruitful results.

Analysis & Research Questions

For the **analysis**, I established research questions to guide through the analysis process. The analysis is divided in a linguistic and an (inter)cultural analysis part drawing on the separation of communicative and intercultural competence throughout the work. The research questions for the respective parts are the following:

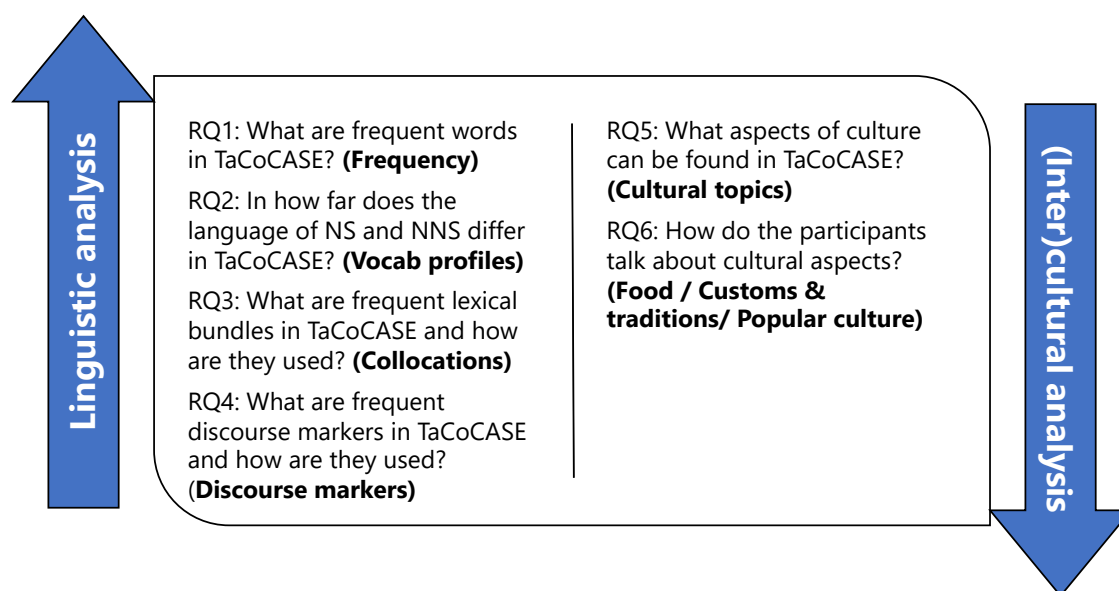


FIGURE 35: RESEARCH QUESTIONS FOR LINGUISTIC AND (INTER)CULTURAL ANALYSIS

As far as the **linguistic analysis** is concerned, research question 1 to 3 are concerned with words and phrases as well as differences in language use between native and non-native speakers and thus draws

on linguistic competence. Research question 4 is concerned with pragmatics and the use of discourse markers and thus draws on pragmatic competence and discourse competence in particular. Linguistic and pragmatic competence are elements of communicative competence in the CEFR and are therefore investigated in the analysis.

As far as the **(inter)cultural analysis** is concerned, I am generally focusing on aspects of Byram's well-established and in teaching widely used model of intercultural communicative competence that is also used as the basis for pluricultural competence in the CEFR (see section 2.1.5.2.3.; see also Byram 1997 and 2021). In research question 5, I draw on the aspect of "knowledge" in Byram's ICC model (*savoir être*, see Byram 1997: 34) focusing on cultural topics that arise in the conversations. In research question 6, I refer to all other factors of intercultural communication, again following Byram's ICC model (see Byram 1997: 34) – skills (*savoir comprendre* and *savoir apprendre/faire*), attitudes (*savoir être*) and education (*savoir s'engager*) (see section 2.1.5.2.3.). RQ 5 is concerned with how the participants talk about cultural aspects in order to find strategies and comprehend how the participants discursively negotiate meaning and mediate cultural concepts.

Linguistic analysis (RQ1-RQ4):

The linguistic analysis combines quantitative and qualitative methods. For RQ 1, RQ3 and RQ4, I used **corpus analysis techniques and methods** such as frequency analysis, n-gram analysis, collocation analysis and concordance line analysis (see section 2.2.1.3.) in order to find frequently used words and patterns and thus look at the language use in the corpus. The corpus analysis software AntConc (Anthony 2023) was mainly used for these corpus searches. RQ1 and RQ3 focus on frequently used words and lexical bundles and in these cases, a quantitative corpus analysis was sufficient. RQ 2 also focuses on quantitative aspects. For this research question, I created vocab profiles of NS and NNS

with Tom Cobb's Vocab profiler¹⁷, which provides information about the words used in comparison to the Academic Word List (AWL) (Coxhead 2000), calculates the type-token ratio of a text as well as the lexical density. These criteria – which will be explained in greater detail in section 4.1.2. – are used to compare the language use between NS and NNS in the corpus TaCoCASE in terms of lexis.

For RQ4, I conducted a **corpus-based discourse analysis**. I extended the quantitative corpus analysis (frequency and collocation analysis) by a qualitative analysis (discourse analysis) in order to find differences in function and use of the discourse markers. I started out with a frequency analysis a first step in order to find frequently used and therefore highly interesting discourse markers (see for example House 2013; Buysse 2012; Collet et al. 2023). After that, I looked at co-occurrences with the help of a collocation analysis. I tagged the discourse markers manually in order to differentiate between discourse marker use and non-discourse marker uses. For the qualitative analysis, I started off with a concordance analysis in order to find patterns of use. In a more close discourse analysis, the discourse markers were analyzed and assigned different functions that they fulfill in the utterances (see House 2013; Buysse 2012; Collet et al. 2023).

(Inter)cultural analysis (RQ5-RQ6):

The (inter)cultural analysis is mainly a qualitative analysis using **discourse and conversation analytical methods**. In order to find cultural aspects, I looked at the conversations and the transcripts and created a conversation profile of each conversation as a first step. The conversation profiles aim to provide a brief overview of the talk. Creating these profiles meant I watched the conversation, looked at the transcript at the same time and I noted down important aspects. The conversation profiles include a summary of the content as well as interesting or remarkable elements of the talk and information on the recording itself. The conversation profiles thus consisted of the following elements:

¹⁷ <https://www.lex tutor.ca/vp/eng/>

- General information (conversation ID, participant information, time of recording)
- Brief content summary (bullet points; including timestamps)
- Salient features: interesting quotes, interesting non-verbal behavior (gestures etc.)
- Recording information (quality, interruptions etc.)
- Keywords

Since RQ 5 is concerned with cultural topics that can be found in TaCoCASE, I looked at the conversation profiles with special regard to the particular cultural topics provided in the topic prompts, namely food, culture and traditions and popular culture. I then extracted examples of these cultural topics from the transcripts and collected them in a separate file (excerpt collection). I identified the cultural topics the participants talked and extracted them from the excerpt collection for the analysis.

In RQ 6, I wanted to investigate how the participants talked about these cultural topics and discursively negotiated meaning and explained cultural concepts. I conducted a corpus-based discourse analysis, with elements of conversation analysis, in which I looked specifically at discourse structure and strategies for mediating (new) concepts. Using DA and CA methods in this intercultural analysis allowed me to analyze the language and the negotiation of (cultural) meanings and concepts. Cohesion and coherence, turn-taking and the use of discourse markers or pauses are aspects of these analyses and help making sense of the utterances and what is meant by the speakers. I investigated the way the speakers structure their utterance, for example when explaining a new concept and developed a discourse structure scheme from the examples I found. This is a very bottom-up approach as I looked at the individual instances first, identified recurring elements and clustered the examples before finding similarities in the discourse structure and labeling the different recurring steps. In order to identify the strategies the participants used to mediate (new) cultural concepts, I conducted a keyword search followed by a discourse and conversation analysis. I created a model describing the strategies the speakers used to explain (new) cultural concepts was based on the individual instances in a bottom-up approach. These kinds of analysis draw on Byram's intercultural communicative competence model (see section 2.1.5.2.3.) and look at both communicative and intercultural elements

at the same time in order to find out about the strategies and the discourse – the “language beyond the sentence” (Yule 2006: 124) – that the speakers use.

Application

Chapter 5 looks at possible **applications of TaCoCASE** and how the corpus and excerpts from the corpus can be used in the EFL classroom. In contrast to the analysis chapter, which is very much influenced by methods and approaches from corpus linguistics, discourse analysis, pragmatics and conversation analysis, this chapter follows a mainly **didactic approach**.

As discussed in sections 2.1.3. and 2.1.4., frameworks and standards like the CEFR, the education standards in Germany and the English syllabi of the federal states in Germany describe competences that should be developed by the learners. Thus, they prescribe what should be taught at school. Many corpus-based pedagogical studies focus on how corpora can be used in the classroom or on one specific aspect like, for example, teaching vocabulary or lexical bundles with corpora. Some recent studies have been published in *The Routledge Handbook of Corpora and English language teaching and learning*, edited by Reka R. Jablonkai and Eniko Csomay (2023), for example studies by Pawel Szudarski on teaching vocabulary, William Crawford on teaching speaking skills, Averil Coxhead on teaching and learning vocabulary in English for specific purposes, Viviana Cortes on lexical bundles in English for Academic purposes, Lynne Flowerdew on English for Academic purposes. These studies, among others, have in common that they focus on aspects of language and how these aspects can be taught with the help of corpora, without linking it directly to the CEFR.

I take a different approach by starting off with the **framework and the competences** that are prescribed by policy-makers and I take that as a **reference point** to see to what extent corpus material can be used to address the same competences. My aim is to demonstrate how well corpora – and in particular TaCoCASE – can be used to establish the competences that are demanded by the frameworks. This is a new approach that other studies only partially or indirectly address by looking at

specific aspects or competences, but they usually do not link it with the CEFR. Thus, I take on a more holistic approach, combining communicative and (inter)cultural competence while at the same time also promoting digital competences by suggesting DDL activities and using corpus methods in the classroom.

Looking at **(inter)cultural competence** from a corpus-based pedagogical point of view is a new approach. Fahey Palma (2023), who is one of the few researchers who looked at promoting intercultural competence with corpora states that “[t]he role of CL in teaching IC [intercultural competence] has, thus, far been limited.” (Fahey Palma 2023: 119). She suggests the use of representative corpora for teaching intercultural skills, which I strongly support. Since digital competences are not mentioned in the CEFR and the use of corpora or DDL is also not suggested, there are no guidelines or objectives to orient to. However, I would argue that these aspects should be integrated in future curricula (see discussion in section 2.1.6.) as they are very important in terms of future skills. DDL has many benefits, above all its learner-centeredness and task-based language learning (TBLL) approach, which facilitates explorative learning and communicative language teaching, both factors of state-of-the art foreign language education (see Gilquin & Granger 2010; see also Decke-Cornill et al. 2015).

The **overall approach** to the use of TaCoCASE in the classroom can be described with the help of the following questions (see figure 36).

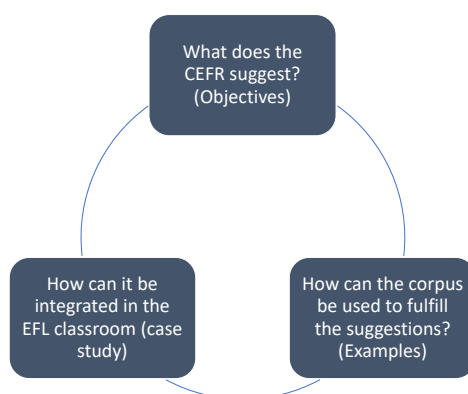


FIGURE 36: TACOCASE APPLICATION PROCESS

The application process starts with the question what the CEFR suggests in terms of objectives for communicative and intercultural competence. Some objectives were selected and examples from TaCoCASE were allocated to the respective objectives. The following figure provides an overview of the selected competences and objectives, which at the same time depicts the structure of this chapter.

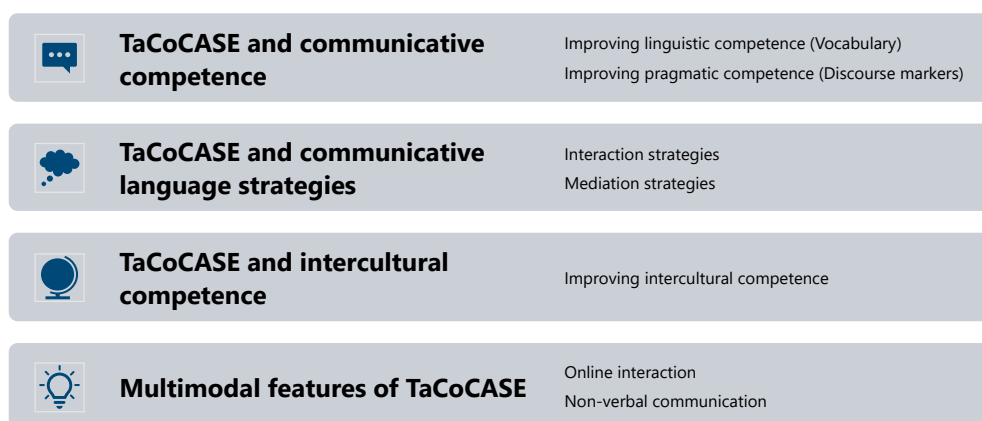


FIGURE 37: OVERVIEW OF SELECTED COMPETENCES AND OBJECTIVES

The elements on the left (see figure 37) are the **competences mentioned in the CEFR** or important elements that should be taught in order to promote communicative and intercultural competence and thus expand language proficiency. The elements on the right, such as improving linguistic competence, interaction strategies etc., are specified by the CEFR in terms of objectives and descriptor scales and key concepts were operationalized for the different levels of proficiency. The competence, the key concepts operationalized and the descriptor scales for the levels of proficiency are described first. In a second step, examples from TaCoCASE, which fit to the objectives and key concepts operationalized, will be presented in order to demonstrate the syllabus fit of TaCoCASE to the CEFR. Due to TaCoCASE's multimodal setup, special features such as online interaction and non-verbal elements can also be subject to analysis. Online interaction and non-verbal elements are mentioned in the CEFR as interaction activities and thus, objectives and key concepts operationalized also exist. The same procedure as illustrated above is therefore applied to the sub-chapter of multimodal features as well.

I chose this approach for **two reasons**: (1) The examples in combination with the suggestions from the CEFR show that TaCoCASE and the corpus data are very well suited for teaching communicative and intercultural competence and thus proof the syllabus fit of TaCoCASE. This is a very strong argument for integrating corpora and CL in general at all stages of EFL teaching and especially in the EFL classroom. (2) The examples could be used in the EFL classroom in order to analyze and comprehend what words, phrases and strategies the speakers used. The learners could use this information and extract these strategies before applying the strategies themselves in situations or practices inside or outside the EFL classroom. Only in combining declarative knowledge and procedural knowledge, the full competence unfolds and only in addressing different learning objectives, the cognitive involvement of the students is higher and with it the learning outcome (see section 2.1.1.).

Learning Objectives

Similar to the research questions that guide through the analysis chapter, I chose to formulate **learning objectives** to guide through the application chapter so that it becomes clear which competences – communicative or intercultural competence – are drawn upon. The learning objectives have two functions: (1) guiding through the individual chapters and formulating the aim what the learners should be able to do after each chapter (“can-do approach”) and (2) making sure that the competences and specific elements of the competences that are demanded by the CEFR are taken into account in teaching. I developed the learning objectives based on the didactic competences mentioned in the CEFR and the education standards (see section 2.1.4. and 2.1.5.). As discussed in section 2.1., it is very important to formulate objectives in teaching in general as they describe what learners should achieve at the end of a certain time span. I follow the common CEFR and education standard approach by formulating the learning objectives with “can do” descriptors, which stress the learner-oriented approach and describes the learners’ abilities (see Elsner 2018). In this case, the learning objectives that I developed describe what the learners should achieve as far as the respective topic is concerned

– ideally at the end of each chapter. The chapters represent different aspects of the competences – communicative and (inter)cultural competence – that I suggest should be promoted with the help of corpus examples. The learning objectives are also developed according to the respective competences and describe elements of the competences demanded by the CEFR.

Sections 5.1. – 5.4. provide **examples from TaCoCASE** that were selected based on the competences demanded in the CEFR. I base my approach of using these examples in the classroom to establish communicative and intercultural competence on well-established cognitive language acquisition theories. Bonnet (2018) summarizes the mental activities, which underly the language acquisition process and the theories behind it, with the following quotation:

From a cognitive point of view, language acquisition is a process that is relatively similar across individuals [...]. It is considered a mental activity of receiving and processing input, negotiating meaning in interactions with interlocutors and generating output that aims at reaching the speakers' communicative goals. The underlying mental activity is seen as a **constructive process** of generating linguistic hypotheses and learners are seen as testing these in interaction (Bonnet 2018: 59; emphasis in original).

I follow this constructive approach and combine behavioristic and cognitive-constructive language acquisition theories and aspects of these approaches, in particular the input theory (Krashen & Terrell 1988) and the interaction theory (Long 1998). Following the input theory, which is known as the natural approach by Krashen and Terrell (1988), learners need comprehensible input in order to be able to establish speaking skills. The examples that are taken from the corpus could serve as comprehensible input for the learners, which the learners can process and which they can convert into intake. The amount of input which can be converted into intake depends on the so-called affective filter: "Given two acquirers with the exact same input, the one with a lower filter will acquire more" (Krashen & Terrell 1988: 38). A low affective filter is desirable so that more input can be processed and converted into output. One factor that influences the affective filter is motivation, which means that the input or in this case the examples from the corpus should be relevant for the learners so that the learners have a high motivation to work with them. The corpus examples are relevant and very close to the learners' real world experience as they are taken from conversations about everyday topics by participants that

are only a little bit older than the learners, which means that the learners could very easily identify themselves with these conversations.

One major criticism of the input theory was that it is mainly based on reception, which means output and especially interaction were not taken into account. I take aspects of the interaction approach into account, which was developed by Michael Long in the 1980s. The interaction approach can be regarded as cognitive-constructivist approach, which means that language acquisition is a mental activity that is driven by interaction and self-organized by the learners (Decke-Cornill 2015: 40). Input and especially comprehensible input play an important role in the interaction theory: “[i]nput must be comprehensible for acquisition to occur”, even though “comprehensible input alone is *insufficient*, particularly with adults and if nativelike proficiency is the goal” (Long 1993: 423; emphasis in original). In contrast to the input theory, the interaction approach states that input can only be made comprehensible through interaction (Long 1993). Interaction and the negotiation of meaning in a linguistic way, which means “to make what they say comprehensible to their interlocutors” (Long 1993: 418) are the key elements of this approach. Long lists several devices or strategies for this negotiation process: “repetitions, confirmations, reformulations, comprehension checks, confirmation checks, clarification requests etc.” (Long 1993: 418). He found out that these strategies occur more frequently in NS-NNS than in NS-NS conversation (Long 1993: 418). The examples that are used in the application are excerpts from the corpus, which consists of interaction between NS-NNS and NS-NS. The examples follow the interaction approach and the observations can be interpreted in view of the interaction approach. However, the learners are not actively involved in the interaction, which means that they can only notice and derive strategies that the participants used, while the speakers were negotiating meaning in interaction with their interlocutors. If the learners take this input, process it and apply it in a context where they are part of the interaction themselves, I would argue that this facilitates the learning process. Learners build upon strategies they have learnt and apply the mechanisms of the interaction theory in their own interaction with other interlocutors. Therefore, it is important to have an application phase after looking at the examples, so that the learners can apply

their knowledge and be actively involved in interaction processes. Without the application phase, the learners would not experience communicative pressure as only looking at the examples and noticing or extracting strategies does not force them to produce output and engage in the language learning process by testing their own hypothesis or applying the interaction strategies. I think it is clear that combining aspects of the input theory and interaction approach as a basis for the use of the examples in the classroom is necessary due to the specific data in the corpus.

In section 5.5., I will go one step further and describe **two case studies**: teaching vocabulary (lexical competence) and explaining cultural concepts (intercultural competence; mediation activity). I decided to present two case studies with DDL activities as more teaching material and hands-on material is usually demanded by teachers and corpus applied pedagogy researchers (see for example Römer 2011; Mukherjee 2002; Le Foll et al. 2021; Pinto et al. 2023). The case studies offer classroom material and demonstrate how TaCoCASE can be used in the EFL classroom, in particular how vocabulary can be taught and how cultural concepts can be explained (mediation). Similar to Pinto et al. 2023 and Le Foll et al. 2021, who created one of the few ready-to-use open educational resources (OER) for DDL with corpus material, I chose to present learning objectives for the lessons along with a concrete lesson plan that includes a step-by-step description of the tasks and content of the lesson, the different phases, interaction forms and methods as well as handouts and material according to didactic standards.

In sum, I follow an interdisciplinary approach, taking corpus linguistics, discourse analysis, pragmatics and conversation analysis as well as applied linguistics and EFL teaching into account in this dissertation. I combined quantitative and qualitative methods from different fields in a mixed-methods approach in order to make up for the complexity of the dataset. This combination of quantitative and qualitative methods allowed an evidence-based as well as close analysis of the data. The application part presents authentic teaching examples and demonstrates the syllabus fit of TaCoCASE. The methodology and the application follows state-of the art teaching methods and provides many examples that could directly be used in the EFL classroom.

4. ANALYSIS: LINGUISTIC AND (INTER)CULTURAL ASPECTS IN TACOCASE

The previous chapter described the database – the corpus TaCoCASE – along with key facts about size, register and compilation of the corpus – as well as the methodology that underlies the analysis and the application. In this chapter, I will analyze the corpus TaCoCASE with regard to **linguistic and (inter)cultural aspects**. The idea is that the results of the linguistic and (inter)cultural analysis build the basis for application in teaching (see chapter 5). The analysis is divided – as throughout this dissertation – in a linguistic and an (inter)cultural part. In the linguistic analysis, the aim is to find out more about actual language in use, which means that I will identify frequent words and patterns in TaCoCASE and analyze how the speakers used them. I will also investigate to what extent the language of NS and NNS is different with regard to lexis, multi-word units and pragmatic markers, such as discourse markers. In the (inter)cultural analysis, the focus lies on cultural aspects and how the speakers mediate and explain (new) cultural concepts. This chapter is guided by research questions for the linguistic and (inter)cultural analysis that will be answered throughout the chapter using a mixed methods approach including quantitative as well as qualitative methods in order to come up with more nuanced findings (see section 3.2.).

4.1. LINGUISTIC ANALYSIS

As far as the **linguistic analysis** is concerned, the aim is to find out more about **actual language in use** in the corpus. This involves mainly quantitative and qualitative corpus methods in order to investigate frequent words and phrases and how they are used by the participants. As the basis for the quantitative analysis, I use corpus methods, including frequency analysis, concordance analysis and collocation as well as n-gram analysis (see O’Keeffe et al. 2007; Hunston 2002; see also section 2.2.1.3.). This approach is typically used in corpus linguistics as a start in order to find frequent words and patterns and find important aspects of language in use, which should be investigated further (Biber et al. 2004: 400). In an interdisciplinary framework, I then add a more qualitative analysis, integrating

elements from discourse analysis and conversation analysis (see section 3.2.). The idea for this qualitative approach is to go beyond frequency and look from a discourse-analytical perspective at how the words are used in context, find patterns and strategies that the speakers used to successfully communicate, organize discourse and negotiate meaning. This involves investigating functions of discourse markers that occurred frequently, finding patterns how the speakers organize discourse when mediating a new concept and a step-by-step analysis how this negotiation process is carried out by the speakers, following the common question in CA: “why that, in what way, right now?” (Seedhouse 2005: 167).

The research questions and the analysis conducted here are based on aspects of **communicative competence** that are important in the application of the corpus in EFL teaching. In turn, the selected aspects of communicative competence are dependent on the corpus methods and tools in order to be able to retrieve findings and examples. As far as improving communicative competence is concerned, I selected two aspects: **Linguistic and pragmatic competence** (discourse competence). Aspects of linguistic competence, such as teaching words and phrases, can be investigated in a quantitative analysis and aspects of discourse competence can be analyzed in a qualitative analysis. Socio-linguistic competence, the third aspect of communicative competence according to the CEFR (CEFR 2001: 108), is not subject to analysis in this case as only one example for a quantitative and one example for a qualitative analysis were chosen.

As far as **linguistic competence** is concerned, teaching vocabulary, including single words and phrases, is one key aspect (CEFR 2001: 108f). Following Michael Lewis’ lexical approach, in which he states that “language consists of grammaticalised lexis, not lexicalised grammar” (Lewis 2002: vi), the focus when teaching vocabulary is on single words and on words in context. The question is what words are worth teaching and how can they be taught in context – and what are important aspects of the context of a word. Identifying frequent words through corpus methods is one way to find single words that are worth teaching. Frequency analysis makes frequently used words visible and “identifies patterns that must be explained” and that “otherwise often go unnoticed by researchers” (Biber et al. 2004: 376).

The idea is that if a word appears frequently in a corpus, it is frequently used by the participants and thus depicts authentic language use, which means it is worth teaching (Römer 2011; see also Biber et al. 2004). Corpus methods provide evidence-based findings for language that is used in the corpus based on frequency. Frequency results from corpus studies “[provide] descriptive facts that require explanation” (Biber et al. 2004: 400). It is thus not possible to interpret the results without looking at the context and explaining what happened, how the words were used, what functions they fulfill in interaction and in discourse and how the speakers made utterances. In my analysis, I therefore combine quantitative corpus analyses such as frequency analysis, collocate analysis and concordance analysis in order to find frequent and useful patterns with a more qualitative approach integrating discourse analytical and conversation analytical methods. The quantitative analysis sheds light on frequent and interesting words and patterns and the qualitative analysis allows interpretations on the functions and use of these words in the negotiation process in meaning making and mediating cultural concepts. Only through the combination of quantitative and qualitative methods, we are able to interpret to what extent language use and negotiation processes are successful in ELF intercultural conversation. When looking at strategies the participants used in the corpus, it is important to see to what extent the strategy was successful or not so that we know what is successful and therefore what is worth teaching and integrating in the EFL classroom.

Based on this mixed methods approach, including elements from corpus linguistics analysis, discourse analysis and conversation analysis, the following research questions have been established:

Research Questions:

- RQ1: What are **frequent words** in TaCoCASE?
- RQ2: In how far does the **language of NS and NNS** differ in TaCoCASE?
- RQ3: What are frequent **lexical bundles** in TaCoCASE and how are they used?
- RQ4: What are frequently used **discourse markers** in TaCoCASE and how are they used?

In section 4.1.1. (RQ1) and section 4.1.3. (RQ3), **frequent words and frequent lexical bundles** will be identified and it will be analyzed how they are used by the participants. Thus, these questions provide

results and information about lexis and phraseology, which are important elements of linguistic competence. Following Michael Lewis' lexical approach and starting out with a frequency analysis as suggested by many applied corpus linguistics (see O'Keeffe et al. 2007; Biber et al. 2004; Hunston 2002), I will begin with looking at frequency of single words and then look at how these words are used in context, including elements of phraseology. This means looking at the periphery of words and extending the scope from single words to collocations (two words used in combination), to multi-word units and lexical bundles (3-grams and 4-grams) (see Römer 2011, Biber et al. 2004). A frequency analysis provides information about frequent words that appear in a corpus and what words are used in this specific register (see Hunston 2002, O'Keeffe et al. 2007). A collocation analysis and an n-gram analysis looks at statistical co-occurrences of words, which provide information on how words are used in combination with each other – either two words (collocations) or more than two words (lexical bundles) – in this case multi-word sequences consisting of 3 and 4 word bundles (McEnery & Hardie 2012; Biber et al. 2004; Hunston 2002). These analysis provide evidence based findings on how words are used in combination and are therefore important to know or teach (Biber et al. 2004; Cortes 2004, Römer 2011). Other lexical bundles consisting of more than 4 words could have been investigated as well, but since 3-grams and 4-grams are most frequently used in spoken register (see Biber et al. 2002), this fits best to the spoken corpus TaCoCASE and to teaching productive language activities (speaking). Following again Michael Lewi's lexical approach and findings from applied corpus linguistics, lexical bundles are found to be easy for learners to retrieve and are used as "discourse building blocks". This means that they are worth teaching and the learners extend productive language activities when using and relying on lexical bundles.

Section 4.1.2. (RQ2) looks at vocab profiles from the individual speakers and **differences in language use between NS and NNS** as far as **lexis** (or vocabulary) is concerned. This research question aims at finding differences or similarities between language use of NS and NNS as far as lexis is concerned in the corpus. NS have for a long time been the model in EFL teaching and NS and NNS are still compared in many research disciplines in terms of language use. I follow the growing criticism of the NS model

(Mukherjee 2002, Byram 1989, Cook 1999, Paikeday 1985 and Piller 2001) and suggest the use of the intercultural speaker as a (new) model in EFL teaching (see discussion in section 2.1.5.2.7.). Based on these statements and based on the ELF intercultural environment that TaCoCASE is set in (see section 3.1.2.2.4.), the language use of NS and NNS in TaCoCASE would be expected not be very different. In fact, the results indicate that there is not much difference between the NS and NNS (see section 4.1.), which underpins these hypotheses.

In order to find differences or similarities in language use, many aspects could be examined, such as lexis, grammar or phonology. As linguistic competence and pragmatic competence will be of interest in the application part in order to establish communicative competence in EFL teaching, I decided to focus on linguistic competence and vocabulary use in this research question. Vocabulary profiles provide information about which words are used, about lexical complexity, lexical variation and density of texts (or conversations) and allow interpretations in terms of language proficiency (see Morris & Cobb 2004; Laufer & Nation 1995; see also section 4.1.2.). Hence, they provide a valid basis for lexical differences, language proficiency and lexical variation in language use among NS and NNS in TaCoCASE. As far as pragmatic competence is concerned, differences in use and functions of discourse markers among NS and NNS have been researched in the past (see House 2013, Buysse 2012, Koops & Lohmann 2016; Collet et al. 2021) and will be investigated in the last research question.

Section 4.1.4. (RQ4) is concerned with **frequency, functions and use of discourse markers**. Using discourse markers contributes to pragmatic competence and discourse competence in particular, which is besides linguistic competence one main aspect of communicative competence. Other factors such as hesitation markers or pauses could as well be interesting in terms of establishing pragmatic competence. I chose to look at discourse markers as markers of pragmatic competence for a number of reasons. First of all, DM appear frequently in the corpus and are thus frequently used by the speakers. They are very frequent in spoken language (Biber et al. 2002) and organize discourse so that coherence is established (see Schiffrin 1987; Yule 2006). Discourse markers are important aspects in establishing pragmatic competence and thus communicative competence in EFL teaching (CEFR 2020).

It is important to teach discourse markers not only on a lexical level, but also on a pragmatic level, focusing on the various functions that the different DM fulfill in discourse and the role they play in establishing coherence in spoken interaction (Kirchhoff 2018: 110). Discourse markers have drawn attention in research in pragmatics as well as corpus linguistics (see Schiffrin 1987, Fraser 1990; House 2013, Buysse 2012, Collet et al. 2021). While corpus linguists typically look at differences in use as far as frequency of discourse markers or phraseology (collocations and n-grams) is concerned, linguists focusing on pragmatics usually look at functions of discourse markers in terms of discourse organization, intersubjectivity or interaction. I combine both approaches and use quantitative (corpus linguistic) methods when looking at frequency and use of discourse markers as well as a discourse analytical pragmatic analysis focusing on the use and functions of discourse markers in terms of discourse organization and the role they play in interaction. Since DM have been well researched in pragmatics and also in ELF research (see for example Buysse 2012, House 2013) and since we conducted analyses of discourse markers in ViMELF already (Collet 2021), which revealed interesting findings, I chose to look at discourse markers in greater detail.

In sum, this chapter and the analysis of the corpus sheds light on the language used in TaCoCASE with regard to lexical and pragmatic elements. It provides evidence-based quantitative and qualitative results that can be used as a basis for the application of TaCoCASE in EFL teaching. The examples that will be presented, analyzed and discussed can serve as examples that could be used in EFL teaching in order to establish communicative competence (see chapter 5).

4.1.1. FREQUENCY ANALYSIS

RQ1: What are frequent words in TaCoCASE?

When analyzing corpora, different searches are conducted that offer “new perspectives on the familiar” (Hunston 2002: 3). In order to find the most frequent words and identify the core vocabulary in TaCoCASE, a frequency analysis was conducted and a word frequency list was created as a first step (see section 2.2.1.3.). Frequency lists display the most frequent words in a corpus and are usually ranked by frequency or alphabetically (McEnery & Hardie 2012; O’Keeffe et al 2007; Hunston 2002). Analyzing a corpus usually involves in a first step a frequency analysis (Evison 2010: 122), which can be used to “identify possible differences between corpora” (Hunston 2002: 5).

Table 20 displays the word frequency list of TaCoCASE generated with AntConc. The first column represents the rank of the word, the second column the type and the third column the raw frequency. Column 4 displays the range, which is the distribution across conversations in the corpus, and column 5 the normed frequency, which is the frequency per million words (pmw), calculated with the following formula (see Brezina 2018):

$$\text{normed frequency} = \frac{\text{raw frequency}}{\text{total number of tokens in the corpus}} \times 1,000,000 \quad (1)$$

TABLE 20: FREQUENCY LIST TACOCASE

Rank	Type	Freq	Range	NormFreq
1	i	5525	15	39,459
2	bi	3418	11	24,411
3	it	3216	15	22,968
4	to	3160	15	22,569
5	sb	3004	10	21,454
6	bo	2964	10	21,169
7	the	2948	15	21,054
8	s	2827	15	20,190
9	you	2624	15	18,740

10	and	2619	15	18,705
11	yeah	2519	15	17,991
12	a	2402	15	17,155
13	like	2308	15	16,484
14	that	2174	15	15,527
15	of	1924	15	13,741
16	t	1775	15	12,677
17	so	1774	15	12,670
18	h	1557	15	11,120
19	in	1408	15	10,056
20	but	1239	15	8,849
21	is	1194	15	8,527
22	uhm	1093	14	7,806
23	u	1085	15	7,749
24	m	952	15	6,799
25	have	927	15	6,621

Table 20 depicts the 25 most frequent different words (types) in TaCoCASE. The most frequent word is *I* and it appeared 5,525 times, which equals about 39,459 pmw, across all 15 conversations. Personal pronouns are very frequent in TaCoCASE as we can find three personal pronouns among the top 10 of the frequency list: *I* (5,525; rank 1), *it* (3,216; rank 3), *you* (2,624; rank 9), which are typically used in spoken discourse as markers of interactivity and which are not very frequently used in written corpora (see O’Keeffe et al 2007: 12). The speaker IDs such as *BI* (3,418; rank 2), *SB* (3,004; rank 5) and *BO* (2,964; rank 6) appear also in the top 10. They are used at the beginning of a turn of the respective speaker and therefore occur very frequently. Frequent verbs are auxiliary verbs such as *BE* (*’s* and *is*) and *HAVE* (927; rank 25), as well as verbs such as *KNOW* (865; rank 28) and *LIKE* (2,308; rank 13). The latter could in fact as well be used as a conjunction or discourse marker, which could only fully determined in a qualitative analysis. Discourse markers such as *yeah* (2,519; rank 11), *so* (1,774; rank 17), *but* (1,239; rank 20) and *uhm* (1093; rank 22) occur also very frequently, which is not surprising

considering that TaCoCASE is a spoken corpus. We find articles such as *the* and *a* (rank 7 and 12), which indicates that a high number of nouns and thus noun phrases is used. *That* occurred also very frequently (rank 14), which generally fulfills two functions: (1) subordinator (following report verbs or it-patterns) and as (2) relative pronoun in relative clauses.

Comparing frequency lists from different corpora provides information about the language used in these corpora and sheds light to words, which appear more frequently than in other corpora of the same genre and register (see Hunston 2002).

In general, it can be stated that the word frequency list from TaCoCASE is quite similar to what O’Keeffe and colleagues (2007: 33-35) found in their spoken corpora from service encounters and friends chatting from LCIE and to the frequency list of CANCODE, which will be explored in great detail later in this section. In these frequency lists, markers of interactivity, such as *I*, *you*, *yeah* (as a response token) occurred very frequently (O’Keeffe 2007: 11). Interactive pronouns and discourse markers are typically found in spoken corpora and they do not appear in written corpora as these features are characteristic for spoken language (O’Keeffe et al 2007: 12). The comparison of TaCoCASE with these two spoken corpora frequency lists suggests that the **frequency lists are very similar** as the same words appear and that **TaCoCASE is very representative for a spoken corpus**.

Another spoken corpus that TaCoCASE can be compared to is ViMELF. The word frequency lists of TaCoCASE and ViMELF are displayed in table 21.

TABLE 21: TACOCASE AND VIMELF - WORD FREQUENCY LISTS

TaCoCASE					ViMELF				
Type	Rank	Freq	Range	NormFreq	Type	Rank	Freq	Range	NormFreq
i	1	5525	15	39459	i	1	5866	20	38646
bi	2	3418	11	24411	sb	2	5805	20	38244
it	3	3216	15	22968	yeah	3	3656	20	24086
to	4	3160	15	22569	it	4	3385	20	22301

sb	5	3004	10	21454	you	5	3182	20	20963
bo	6	2964	10	21169	to	6	3056	20	20133
the	7	2948	15	21054	s	7	3032	20	19975
s	8	2827	15	20190	h	8	2965	20	19534
you	9	2624	15	18740	and	9	2947	20	19415
and	10	2619	15	18705	the	10	2642	20	17406
yeah	11	2519	15	17991	a	11	2422	20	15956
a	12	2402	15	17155	t	12	2322	20	15298
like	13	2308	15	16484	he	13	2086	19	13743
that	14	2174	15	1,527	that	14	1960	20	12913
of	15	1924	15	13741	so	15	1764	20	11621
t	16	1775	15	12677	nods	16	1741	14	11470
so	17	1774	15	12670	in	17	1730	20	11397
h	18	1557	15	11120	uh	18	1676	20	11042
in	19	1408	15	10056	of	19	1586	20	10449
but	20	1239	15	8849	like	20	1507	20	9928
is	21	1194	15	8527	uhm	21	1450	20	9553
uhm	22	1093	14	7806	fl	22	1423	5	9375
u	23	1085	15	7749	but	23	1392	20	9171
m	24	952	15	6799	st	24	1374	12	9052
have	25	927	15	6621	have	25	1343	20	8848
head	26	877	13	6263	head	26	1309	14	8624
hm	26	877	15	6263	u	27	1304	20	8591
know	28	865	15	6178	sf	28	1257	5	8281
they	29	845	15	6035	we	29	1020	20	6720
just	30	831	15	5935	is	30	1014	20	6680

22 out of the top 30 words in TaCoCASE and ViMELF are the same for both corpora, even though the normed frequency and the ranks of the respective words vary. *It*, for example, is used in both corpora, but slightly more frequently in TaCoCASE (22968 pmw), where it is the third most frequent word in the corpus, compared to ViMELF (22301 pmw), where it is the fourth most frequent word. The differences, however, are very small and not salient. The fact that more than 2/3 of the words are the same for both corpora, is more important. It can be stated that in terms of frequent words, the two corpora do

not differ very much and are very similar. This is not very surprising as TaCoCASE and ViMELF are both spoken corpora and they are both compiled in a similar manner. It does not make a huge difference whether the corpus consists of NS and NNS or only of NNS data, which is a sign that both corpora can be regarded as corpora in an ELF environment.

Only 8 words in TaCoCASE and ViMELF are different (the different words are highlighted in table 21). Even the most frequently used word *I*, the personal pronoun, is the same for both corpora with almost the same frequency (normed frequency in TaCoCASE is 39,459 pmw, compared to 38,646 pmw in ViMELF). Apart from that, we find similarities in both corpora, such as personal pronouns (*I, you, it*), auxiliary verbs (*have, BE*), discourse markers (*yeah, so*) and hesitation markers (*uhm*), prepositions (*in*) and conjunctions (*and, but*). The personal pronouns (*I* and *you*) as well as discourse markers are usually frequently found in spoken corpora as they serve important aspects of interaction, such as responding or backchannelling (O’Keeffe 2012: 34). Differences arise from the different speaker IDs in TaCoCASE (BI, BO, SB) and in ViMELF (SB, HE, FL, ST, SF), which is due to the fact that we have different speakers in the two corpora and the speaker IDs are at the beginning of a turn to indicate which speaker’s turn it is.

It is interesting to see that *nods* as a non-verbal element signaling approval and backchannelling is in the top 30 list for ViMELF, whereas markers of non-verbal communication are not found among the top 30 in TaCoCASE (*nods* appears on rank 55 in TaCoCASE). There seems to be a difference in non-verbal communication among participants from TaCoCASE and ViMELF as far as frequency is concerned. In this particular case with *nods*, the difference seems to be in backchanneling behavior to signal approval. Meierkord (2000) for example found that supportive backchannelling is – besides face-saving acts – one key principles in ELF and intercultural communication. She found instances of backchannelling that were expressed verbally as well as non-verbally (in this case through supportive laughter). Research from a multimodal perspective on backchannelling find instances of verbal responses – such as *yeah* or *okay* – as well as non-verbal responses – such as facial expressions, gestures and nods (Tolins & Tree 2014; see also Bertrand et al. 2007). In ELF conversations,

backchannelling was found to be very frequently used non-verbally (Bjørge 2010) or with minimal responses such as *yes/yeah* or *okay* (Mauranen 2006: 147). Bjørge (2010) investigated listener behavior and backchannelling in particular and found that nods were the most frequent backchannel in her study of video recordings of simulated (business) negotiations. This is true for ViMELF as well. In terms of verbal backchannelling, Bjørge found *yes/yeah*, *mhm* and *okay* to be very frequent (Bjørge 2010, see also Mauranen 2006).

Following previous research as stated above (see for example Meierkord 2000, Bjørge 2010 and Mauranen 2006) and taking the ELF intercultural environment in TaCoCASE and VMELF into account, there should be no difference between TaCoCASE and ViMELF. However, in ViMELF, it seems that verbal as well as non-verbal backchannelling is more frequent. Nods are more frequent in ViMELF (rank 16) than in TaCoCASE (rank 55), as previously mentioned. Verbal backchannelling with *yeah* is also more frequent in ViMELF (24086 pmw; rank 3) compared to TaCoCASE (17991 pmw; rank 11). Differences in frequency of non-verbal communication and nods in particular in TaCoCASE and ViMELF could either be a sign for different uses, which means that – in this case – backchannelling is more frequently used by the speakers in ViMELF. The difference in frequency of non-verbal communication could as well arise from different transcribers and differences in perception of salience of the non-verbal elements – even though this factor is quite small as the transcripts fulfill a qualitative high standard with many correction phases. It could also be a sign of cultural differences in ViMELF and TaCoCASE as backchannelling and non-verbal behavior is also very frequently used in negotiating meaning in creating rapport in the ViMELF conversations, for example to make up for lexical gaps (see Brunner 2021; Antonello 2022). It seems that the setting in TaCoCASE, the intercultural ELF environment, as well as the fact that there is not much lexical difference between NS and NNS (see results in section 4.1.2.), could have an influence on the lower frequency of non-verbal elements in TaCoCASE compared to ViMELF. A closer qualitative analysis would have to be conducted in order to find out more about the differences of non-verbal communication and backchannelling with regard to

the speakers' possible intentions. This is a very interesting aspect that should be investigated further, but it would go beyond the scope of this work to do so.

The first person plural personal pronoun *we* appears only in the top 30 list for ViMELF (rank 32 in TaCoCASE) whereas the third person plural personal pronoun *they* appears only in the TaCoCASE top 30 list (rank 36 in ViMELF). Referring to “us” versus “they” is a means to express cultural identity and social representation (Howarth 2002; Moscovici 1961). It is a very common phenomenon in intercultural communication that speakers position themselves towards a specific culture and negotiate identities discursively and socially in conversation (see Edwards 2009, Baker 2015 and Swann & Bosson 2008). The speakers take on “different identities at different times” in the conversations (Hall 1992: 277) and thus use either an inclusive *we* or referring to other people in using *they*. A more qualitative analysis would have to be conducted to look at *we* in more detail, whether it is used as inclusive *we*, referring to the two speakers in the conversation, or an exclusive *we*, referring to the speaker as a member of another group or culture (see Harwood 2005). These two features – non-verbal elements and plural personal pronouns – play an important role in meaning making and the creation of common ground as part of the ELF intercultural setting and the intercultural competence of the speakers (see section 2.1.5.2.).

Comparing the frequency lists of TaCoCASE and ViMELF to another spoken corpus, CANCODE, a five million corpus of spoken interaction compiled by McCarthy and Carter in the 1990s, reveals similarities in both corpora. 19 out of the 30 most frequent words in TaCoCASE and ViMELF appear as well in the top 30 list in CANCODE (see O’Keeffe 2007: 35). The frequencies of the individual words are even quite similar, for example *I* appears in CANCODE 30197 pmw (overall frequency is 150989; see O’Keeffe 2007: 35) and in TaCoCASE 39459 pmw and in ViMELF 38646 pmw. This shows that TaCoCASE and ViMELF are very comparable to other spoken corpora and thus are representative of the spoken register. The frequency lists show many markers of interaction and intersubjectivity, such as personal pronouns, discourse markers and hesitation markers. These markers are very commonly used in spoken interaction and in spoken language in general, which makes them important aspects for

establishing linguistic competence in EFL teaching. In general, the most frequent words in TaCoCASE are very representative of the spoken register – as the comparison with other spoken corpora showed – and these words are worth teaching in order to establish communicative competence and linguistic competence in particular.

4.1.2. LEXICAL ANALYSIS

RQ2: In how far does the language of NS and NNS differ in TaCoCASE?

Differences in language use among NS and NNS have been subject to analysis in many different research areas. Especially in EFL teaching, native speakers and their language use have for a long time been taken as model or reference for learners (see section 2.1.5.2.7). With growing criticism about the native speaker model (see for example Byram 1997, Cook 1999, Paikeday 1985 and Piller 2001), ELF-aware teaching and teaching intercultural communication in a more hybrid and transcultural sense, the distinction between NS and NNS declines and the intercultural speaker has become an alternative and in my opinion more suitable model (see discussion in section 2.1.5.2.7.). The intercultural speaker fits very well to the two corpora ViMELF and TaCoCASE as they both depict intercultural encounters and the setting can be described as intercultural ELF setting (see section 3.1.2.2.4.). The participants in ViMELF and TaCoCASE can be regarded as intercultural speakers who use English as lingua franca in an intercultural setting, where meaning is negotiated – linguistically and culturally – in interaction. This section is intended to investigate to what extent the language of NS and NNS actually differs in general and in the corpus TaCoCASE in particular. Taking the ELF intercultural environment of TaCoCASE into account, I would expect that the language use of NS and NNS does not differ very much as the speakers can be regarded as intercultural speakers.

In order to find **differences (or similarities) in the language use by NS and NNS**, several criteria could be taken into account, such as lexis, grammar or phonetics. As linguistic and pragmatic competence play an important role as far as learners' communicative competence is concerned (see CEFR 2020)

and these are the aspects that will be of interest in the application chapter, the focus will be on different uses of linguistic and pragmatic features between NS and NNS. Pragmatic differences will be addressed and discussed in section 4.1.4 with regard to discourse marker use among NS and NNS. In terms of linguistic differences, the focus will be on **lexis** and specifically the **vocabulary and words the speakers use**.

For this **lexical analysis**, I used Tom Cobb's Vocab profiler¹⁸ and created a **vocab profile** for the speakers. Vocabulary profiles are used to examine linguistic differences among NS and NNS which is usually interpreted in terms of language proficiency (Morris & Cobb 2004: 77). The idea is that vocabulary proficiency provide insights into linguistic and academic ability (see Laufer & Nation 1995; Morris & Cobb 2004). Morris and Cobb (2004) for example looked at vocabulary profiles and the correlation of academic success and performance in student entry essays. They found significant correlation between vocab profiles and grades, which means that vocab profiles are a good predictor for academic success. In a similar manner, Vo (2019) looked at lexical features in NNS academic writing in a corpus of English placement tests and found a correlation between academic success and linguistic proficiency. Leńko-Szymańska (2015) examined proficiency among NNS across different education systems. She found that the lower the amount of K1 words and the higher the proportion of AWL words, the higher the proficiency level in the CEFR. These findings suggest a correlation between vocabulary profiles and linguistic ability as well as language proficiency. It is to mention here that vocabulary profiles and proficiency have been studied with regard to written texts and only very rarely with spoken discourse. Christoffersen (2017) is an exception as she investigated oral proficiency of NNS, using quantitative methods (TTR) and NS ratings.

Vocabulary profiles provide the following information, which can be extracted with the help of Tom Cobb's vocab profiler:

¹⁸ <https://www.lex tutor.ca/vp/eng/>

- **Number of K1 words**; the first 1000 most frequent words in English based on findings from the BNC and COCA
- **Number of K2 words**; second 1000 most frequent words in English (number 1001 until 2000).
- **Number of AWL words**; words which can be found on the Academic Words List
- **Number of off-list words**; very rarely used words and not on the AWL
- **Lexical density**; the ratio of content words per total words calculated by the following formula:

$$\text{lexical density} = \frac{\text{content words}}{\text{total number of words}} \quad (2)$$

- **Type-token ratio (TTR)** can be calculated according to the following formula (Brezina 2018: 57):

$$TTR = \frac{\text{number of types}}{\text{number of tokens}} \quad (3)$$

The first four criteria – **K1, K2, AWL and off-list words** – provide information of the different words the speakers used and thus about complexity of the language they used. If a text (or in this case a transcript) has a high number of K1 words, it means that the text is very comprehensible as it consists to a large part out of the 1000 most frequent words in English. A high number of K1 words is expected as TaCoCASE belongs to the spoken register and spoken language is usually not as complex as written language. If a speaker uses mostly K1 words and only a few AWL words, the academic proficiency would not be very high (see Leńko-Szymańska 2015).

The **lexical density** describes – as the name implies – how lexical dense or varied a text is. It was introduced by Ure (1971) and describes the ratio of content words (lexical items) and the total number of words or tokens in a text. The result is always a number between 0 and 1 as calculated by equation (2), which means the number of content words divided by the total number of words in the text (or transcript). The higher the ratio is, the more content words a text entails and thus, the more varied the text is (Brezina 2018: 57). Lexical density is a measure for the complexity of a text, typically used with written texts (Halliday 1985: 62). As far as register is concerned, academic writing has a higher lexical density than informal texts (see Biber 2006).

The **TTR** is the “number of different words (types) in a text divided by the total number of words (tokens)” and it is a measure for lexical diversity or variation in a text (Cobb et al. 2015: 192). The TTR is always a value between 0 and 1. If the ratio is closer to 1, the number of types and tokens is closer to each other, which means that there is less repetition and more lexical variation. The larger the number of types, the more varied is a text (Brezina 2018: 57). TTR has sometimes been criticized as it is dependent of text length (see for example Brezina 2018). If texts differ too much in total words, TTR does not provide comparable results. Sometimes this deficit is made up for in using Standard Type/Token ratio (STTR) (Brezina 2018: 57). As the texts in this analysis are quite equal in size, the TTR is sufficient for the purpose of this study.

Lexical density and TTR are both used as a measure of **lexical variation** – however, they differ in the terms what they measure (see Johansson 2008). The denominator of fraction (2) and (3) are the same – number of tokens and number of words is the same in this case. This means that the fractions differ only in terms of the numerator and thus lexical density is dependent on content words and TTR is dependent on types. Content words are not the same as types, which can be any kind of different words in a corpus, which means that the numbers are not the same and the ratios vary. All these values and criteria provide information about **lexical variation and preferences** among NS and NNS and can be used to compare vocabulary use and lexis among NS and NNS.

Vocab Profiles

When comparing the **vocab profiles** for TaCoCASE and ViMELF, there seems to be no difference at all between the two corpora (see table 22) as far as lexis is concerned.

TABLE 22: COMPARISON OF VOCAB PROFILES FOR TACOCASE AND VIMELF

mean	K1 words	K2 words	AWL words	Off List words	Lex density	TTR
TaCoCASE	81.58	3.62	1.51	13.32	0.48	0.20
ViMELF	80.32	3.49	1.29	14.90	0.48	0.19

The numbers depicted in table 22 for TaCoCASE and ViMELF are very similar with only marginal differences. There is a small difference in K1 words between TaCoCASE and ViMELF (difference is 1.26) and an even smaller differences for K2 words, AWL words and off list words. There is no difference in lexical density and almost no difference (0.01) in TTR between TaCoCASE and ViMELF. Thus, the two corpora seem to be very similar in terms of language use, which underpins the fact that both corpora display the same language and are representative for spoken discourse and in particular intercultural communication in an ELF setting. However, it is worth looking at the individual speaker groups for a more detailed analysis.

The following tables (tables 23 and 24) display the **vocab profiles** for the speaker groups in TaCoCASE and ViMELF and the mean values for K1 words, K2 words, AWL words and Off list words.

TABLE 23: VOCAB PROFILE TACOCASE

Speaker group (mean)	K1 words	K2 words	AWL words	Off List words
BI	80.07	3.75	1.56	14.71
BO	80.91	3.60	1.73	13.77
SB	83.76	3.53	1.23	11.48

TABLE 24: VOCAB PROFILE VIMELF

Speaker group (mean)	K1 words	K2 words	AWL words	Off List words
FL	79.32	3.69	1.03	15.96
HE	80.63	3.02	1.27	15.08
SB	80.28	3.60	1.19	14.93
ST	81.11	3.43	1.85	13.60

The numbers for all speaker groups for K1 words are very close at around 80%. The highest number of K1 words is reached by the German speakers in TaCoCASE (83.76%), whereas the lowest number of K1 words is reached by the Italian speakers (79.32 %). The Finnish speakers use the fewest K2 words (3.02%) and the British speakers use the largest amount of K2 words (3.75%). The Spanish speakers use AWL words most frequently, followed by the native speakers (BI=1.56%, BO=1.73%). The Italian speakers use AWL words least frequently (1.03%). All in all, it can be stated that the vocab profiles are

very similar for NS and NNS in general. Some speaker groups may have slightly different values for one of the categories, but overall, the vocab profiles look very similar.

Morris & Cobb (2004) investigated written student entry essays and their lexical density. They state that an ideal vocab profile is reached if the number of K1 words is smaller than 85% and the number of AWL words is larger than 5% (Morris & Cobb 2004: 83). In their study, 62% of the NS and only 29% of the NNS reached ideal K1 lower than 85% and 64% of NS and 46 % of NNS reached AWL larger than 5%. When comparing the results to TaCoCASE and ViMELF, it can be stated that all speaker groups in TaCoCASE and ViMELF have an ideal vocab profile in terms of K1 word values, but all speaker groups do not reach the ideal profile with regard to AWL word numbers as they are below 5%. Since the numbers are so close for TaCoCASE and ViMELF, I would argue that the main difference lies in the register. Morris and Cobb looked at written essays in an academic context (student entry essays) whereas TaCoCASE and ViMELF represent spoken informal interaction between students. Spoken interaction, as carried out in TaCoCASE and ViMELF, is characterized by a high number of content and most frequent basic vocabulary (K1 words) and a low number of academic vocabulary (AWL words). It can be summarized that the results in terms of K1 words are in line with Morris and Cobb, but the AWL words differ, which is mainly due to the register.

It is interesting to see that there is a difference in the German speaker's K1 words use in TaCoCASE compared to ViMELF. The German speakers in TaCoCASE use 83.76% K1 words and 11.48% Off list words, whereas the German speakers in ViMELF use 80.28% K1 words and 14.93% Off list words. The number of K2 words and AWL words is quite similar. One possible interpretation could be that the German speakers adapt their language to their interlocutors. In TaCoCASE, the German speakers, who are non-native English speakers, talk to native English speakers and in ViMELF, the German speakers talk to other non-native speakers. When talking to NS, German speakers seem to use more K1 words and when talking to NNS, they seem to use fewer K1 words. Their lexical choice (and hence their proficiency) varies according to their interlocutors. In fact, it has been stated in research that ELF speakers adapt to their interlocutors and the speakers' communicative needs, which means that they

use communicative resources differently and correctness and appropriateness varies in different conversations (see for example Seidlhofer & Widdowson 2017, Baker 2015, Jenkins 2015b, Widdowson 2012). It may be the case that the German speakers identify as learners of English in the conversations with the NS and that they identify as intercultural speakers in conversations with other NNS. When identifying as learners, the German speakers probably emphasize their – according to their learner's mindset – “less” proficient language use which has an effect on vocabulary choice. It seems that they use more K1 words and less variation in order to feel more comfortable. When identifying as intercultural speakers, it seems that language proficiency does not play an important role in an ELF intercultural environment and that successful communication is more important (see Jenkins 2015a, Mauranen 2012, Seidlhofer & Widdowson 2017). However, it is very hard to tell just from looking at the frequency counts and without conducting a more qualitative analysis to find out more about the participants' motivation for vocabulary use or their feeling of identity in the interactions. It would be interesting to investigate this in future research.

When looking at the vocab profiles of the individual speakers only in TaCoCASE, there is only one NS that has a number of K1 words higher than 85%, which is BI33 (K1 = 86.74%). As far as NNS are concerned, there are three NNS with a K1 word number higher than 85%: SB124 (K1= 86.68%), SB147 (K1=87.27%) and SB149 in the conversation with BO03 (K1=85.72%). The majority of NS and NNS still uses less than 85% K1 words and reaches an ideal vocab profile (Morris & Cobb 2004).

It is surprising that we find differences in use of K1 words for one speaker, SB149, who held two conversations. In conversation 03SB149BO03, SB149 uses 85.72% K1 words and in conversation 07SB149BI31, SB149 uses 81.07% K1 words. It seems that the same speaker can have an ideal vocab profile in one conversation and in another conversation, the profile would not be ideal in terms of K1 words. This means that the use of K1 words varies across conversations and it is not a stable factor for speakers. This would underpin the hypothesis that speakers adapt the language to their interlocutors and the interaction's setting as previously stated. TaCoCASE and ViMELF are set in an ELF and intercultural environment, where speakers negotiate meaning and identities in interaction with their

interlocutor. Thus, speakers show different lexical preferences in different intercultural interactions and they change preferences in different conversations. It might be interesting to investigate one speakers' vocab profiles across different conversations and examine lexical choice in future research. A qualitative analysis with speaker interviews would have to be conducted in order to investigate this further.

In general, it can be stated that the **vocab profiles** of the speaker groups in TaCoCASE and in ViMELF indicate that there is **not much difference between language use** – in this case lexis – of NS and NNS as the numbers are quite similar. Thus, there is no big difference between native and non-native speakers language use considering the vocab profiles.

Lexical Density

Another factor, which could reveal differences in language use and especially in lexical complexity is the **lexical density**. Lexical density is measured according to equation (2):

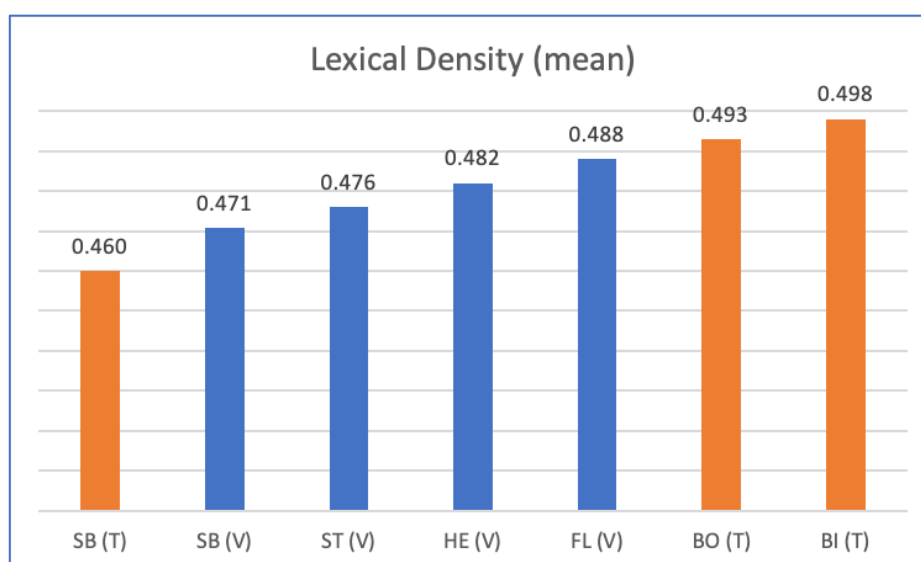
$$\text{lexical density} = \frac{\text{content words}}{\text{total number of words}} \quad (2)$$

A low lexical density indicates that fewer content words are used in comparison to the total words and thus it is a measure for lexical variation and complexity (see Ure 1971; Halliday 1985). If fewer content words are used, more non-content words such as conjunctions or hedges are used and the texts are less lexically rich and thus not so different and varied in terms of lexis. In this case, a higher lexical density indicates a higher use of content words, which in turn leads to more complexity and more variation in a text. Several studies investigated the correlation between lexical density and lexical and linguistic proficiency and found a strong correlation between the two (Nasseri & Thompson 2021; see Lu 2012; Brezina & Palotti 2019, Kim 2014; Linnarud 1975). The higher the lexical density, the higher the lexical or linguistic proficiency (see Kim 2014; Linnarud 1975). As far as the difference between NS

and NNS is concerned, the picture is not very clear. Some studies suggest that NS show more complexity in their language use and thus a higher proficiency than NNS (see Nassari & Thompson 2021; Gonzalez 2013). Other studies did not find differences between NS and NNS as far as complexity and lexical density is concerned (Brezina & Palotti 2019; Pietilä 2015).

Figure 38 displays the mean **lexical density** for the different speaker groups in ViMELF (V) – displayed in blue bars – and TaCoCASE (T) – displayed in orange bars.

FIGURE 38: LEXICAL DENSITY (MEAN)



The native speakers in TaCoCASE demonstrate the highest lexical density, which is in both cases above 0.49 (BO= 0.493 and BI= 0.498). In the case of the British speakers, the lexical density and thus the relation between content words towards and total words in the corpus is almost 1:2, which means that almost every second word is a content word. The results indicate that non-native speakers have a slightly lower lexical density in their texts and thus in their conversations compared to the native speakers. The German speakers from TaCoCASE show the lowest lexical density (0.46). As described above, lexical density is a measure for lexical variation. Since native speakers have a slightly higher lexical density, they use more content words compared to the non-native speakers and thus native speakers show greater lexical variation in their conversations. The numbers for the NS and NNS,

however, are still very close and the lexical density is quite similar (difference = 0.038). This means that even though we see a slight tendency for NS to have a higher lexical density compared to NNS, in general, the lexical density is quite similar for NS and NNS.

The native speakers in TaCoCASE show – in general – a higher lexical density, which means that they show a higher lexical complexity (Nasseri & Thompson 2021; Lu 2012), which results in a higher language proficiency (see Brezina & Palotti 2019; Kim 2014; Linnarud 1975). This supports findings that NS show higher lexical complexity, variation and thus a higher language proficiency compared to NNS (see Nasseri & Thompson 2021; Gonzalez 2013). However, the difference is quite small and not salient. This underpins the fact that the difference between NS and NNS in terms of lexical density is not very clear as there are studies who found significant differences and others who did not. What seems to be clear though is a correlation between proficiency and lexical density (see Nasseri & Thompson 2021; Lu 2012; Brezina & Palotti 2019, Kim 2014; Linnarud 1975). In this case, the NS and NNS in TaCoCASE are quite equally proficient as there is no big difference between NS and NNS in terms of lexical density.

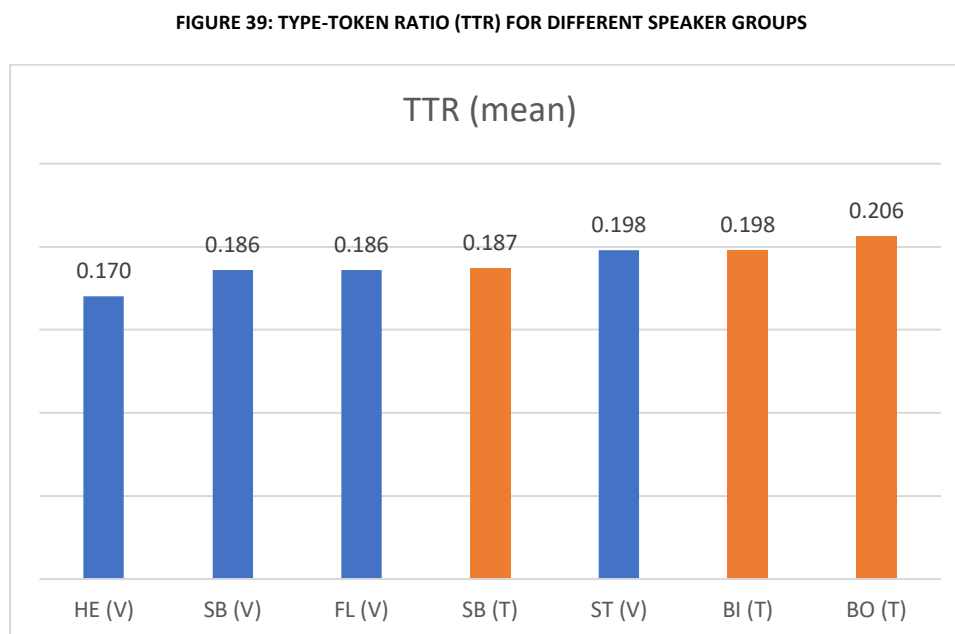
Type-token ratio

Another factor for variation or lexical diversity in a text is the **type-token ratio**. The TTR is calculated with equation (3), which means it is the ratio between types and tokens in a text (Brezina 2018).

$$TTR = \frac{\text{number of types}}{\text{number of tokens}} \quad (3)$$

Types are different words in a corpus and tokens is every single word in a corpus. If a text has many types, this means that there are many different words in a corpus. Thus, the larger the number of the numerator of equation (3), the more different types (different words) are used and the more varied and complex is a text (Brezina 2018: 57). If fewer types are used, words are repeated and thus the text is not varied too much.

Figure 39 displays the mean values for the **TTR of the different speaker groups** in TaCoCASE (bars in orange) and ViMELF (bars in blue).



As figure 39 shows, native speakers have the highest TTR in their texts and the TTR for almost all non-native speakers is below the native speakers' number, except for the Spanish speakers, which have an equal TTR as the British speakers (both 0.198). Finnish speakers have the lowest TTR with 0.17, followed by the German speakers in ViMELF and the Italian speakers (both 0.186). The TTR between the German speakers in ViMELF and in TaCoCASE is not very different (difference = 0.001) as the German speakers in ViMELF have a TTR of 0.186 and the German speakers in TaCoCASE have a TTR of 0.187. Similarly to the results from lexical density, it can be stated that the native speakers – in general – have a higher TTR. The results indicate that NS use more types in a text and thus more different words. Using more types and more different words is a sign of less repetition and more variation (Brezina 2018: 57). It can thus be stated that the NS in TaCoCASE show more lexical variation in their texts. However, the numbers are still very close (difference = 0.03) and it can be stated that the TTR for NS and NNS is very similar, which means that there is no big difference in TTR between NS and

NNS and thus no big difference in lexical variation between NS and NNS. This is in line with the findings from lexical density in TaCoCASE and ViMELF.

The numbers for TTR are quite in line with what Kettunen (2014) found, when investigating EU constitution texts and random sentence data from newspaper and web material (what they called the Leipzig corpus in the study). Kettunen found a TTR between 0.1 (for EU constitution texts) and 0.23 (for then Leipzig corpus). In TaCoCASE, the mean TTR for the speaker groups lies between 0.17 and 0.2, which is the same range compared to Kettunen (2014) even though the register is again different (Kettunen uses written texts and TaCoCASE entails spoken interaction).

In general, it can be concluded that there is not a big difference in vocab profiles and thus in lexis for the speaker groups in TaCoCASE and ViMELF. The language use in terms of lexis of non-native speakers does not differ very much from the language use of the native speakers. Slight differences in vocab profiles for German speakers and for one individual speaker who has two different vocab profiles in two different conversations indicate that vocab profiles might not be dependent so much on speakers but on conversation. This means that speakers adapt to the interlocutors and use different words or different vocabulary in different conversations, a phenomenon that has been found in ELF research (see for example Long 1996, Jenkins 2015b, Mauranen 2012, Seidlhofer & Widdowson 2017). However, this would have to be investigated further more qualitatively in order to make reliable assumptions. Slight difference can also be found in terms of lexical density and in type-token ratio. NS conversations show a slightly higher lexical density and a higher type-token ratio, which indicates that the native speakers use slightly more variation in their utterances compared to the non-native speakers.

It can be concluded that **no big differences between NS and NNS** exist in terms of lexical choice based on the three measures (vocab profiles, lexical density, TTR). This is a sign that the speakers in TaCoCASE and in ViMELF have a very **similar language use** in terms of lexis, which means that the difference between NS and NNS, which is usually promoted in EFL teaching, is in fact not there. NS and NNS in TaCoCASE and ViMELF show similar language use and therefore, it could be argued that nativeness is

not a good criterion to distinguish the speakers. The fact that NS and NNS do not show big differences in language use – in this case lexis – provides proof for the ELF and intercultural environment in TaCoCASE. In an **ELF intercultural environment**, ELF speakers negotiate and mediate language and cultural concepts in interaction with interlocutors. Speakers in TaCoCASE and ViMELF show that they belong to the same group of speakers: intercultural ELF speakers. The differentiation between NS and NNS is thus not suitable any more in this ELF intercultural setting.

4.1.3. COLLOCATIONS AND N-GRAMS

RQ3: What are frequent lexical bundles in TaCoCASE and how are they used?

There are many terms and expressions for **strings or sequences of words** such as multi-word-sequences, lexical bundles, clusters or n-grams, which basically mean the same, but are used in different contexts. Firth's quote "you shall know a word by the company it keeps" (1957: 11) is very famous and widely used when looking at strings of words or words that are used in combination with each other. Form and meaning of a word go hand in hand and meaning of a word can only be constructed through looking at the context in which it is used.

Collocations are two (or more) words that frequently occur together and that gained much attention in corpus linguistics (Yule 2006: 108). A collocation defines "the relationship that a lexical item has with items that appear with greater than random probability in its textual context" (Hoey 1991: 6) or simply put, it describes "a relationship between words that frequently occur together" (Yule 2006: 239). When looking at collocations from a corpus linguistics perspective, frequency and statistical information are important and the co-occurrence of words is accompanied by a high probability, which means that there is a strong combination and a high probability that the words occur together and their co-occurrence is not by chance.

Collocations can be extended to frequently co-occurring multi-word sequences that are usually referred to as **lexical bundles**. Lexical bundles are defined as “extended collocations, sequences of three or more words that statistically co-occur in a register” (Cortes 2004: 400, see also Biber et al. 1999). According to Biber and Barbieri (2007: 269), lexical bundles have three features that differentiates them from other formulaic expressions: (1) they are “extremely common”, (2) “not idiomatic in meaning” and (3) “not perceptually salient”. Lexical bundles are “frequently re-used, and therefore become ‘prefabricated chunks’ that speakers and writers can easily retrieve from their memory and use again as text building blocks” (Biber et al. 2002: 443). In terms of frequency, Biber and colleagues state that a lexical bundle must occur at least 10 times per million words and they have to appear over a number of texts, i.e. occurring only in one text sample would not be a lexical bundle according to this definition. Cortes (2004: 404) sets the frequency cut-off point even higher and only counts occurrences of 20 times per million words or more. In a similar manner, Hyland defines lexical bundles as “building blocks for coherent discourse” (Hyland 2008: 6), stressing their discourse-building function. Lexical bundles do not necessarily have to carry meaning, but are found to frequently co-occur in language. Only 15% of lexical bundles in conversation count as semantically fine or complete, which means that most expressions are incomplete (Biber et al. 2004).

According to Biber (2009), there are two types of **multi-word-sequences**: (1) Multi-word lexical collocations, which are combination of content words and (2) multi-word formulaic sequences which consist of function words and content words (Biber 2009: 290). The difference in these multi-word sequences lies in their degree of fixedness. Typically, strings of words are depicted on a continuum in terms of function and meaning, starting with free combinations over bound combinations (collocations) to frozen combinations (idioms). In a similar manner, Ellis (2008) and Holme (2009) use a construction schema: Filled (idioms), partially-filled and unfilled with an abstract schema (see also Liu 2012; see section 5.1.1.). The idea is that in free combinations, words can more easily be substituted than in fixed combinations, where the combination is fixed and the words have to be in the correct sequence so that the meaning does not change. Multi-word sequences are especially useful

in EFL teaching, where they are usually referred to as lexical bundles or simply (lexical) chunks, which will be demonstrated in the application chapter (see section 5.1.1.).

It is important to note that lexical bundles or collocations should not be confused with **idioms**. All of them are multi-word sequences of words, but the difference is that the meaning of idioms cannot be predicted by the meaning of the individual words and they usually have to be interpreted in a metaphorical way (Biber et al. 2002: 447). Idioms are fixed expressions, which means that words could not be exchanged without changing the meaning of the expression. This draws upon Sinclair's idiom principle, in which he states that "[a] language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments" (Sinclair 1991: 110).

A more technical way of expressing strings of words or multi-word sequences, especially in corpus linguistics, is using the term n-grams, where n is the number of words that the multi-word sequence consists of (see Biber et al. 1999). This goes back to the n-gram analysis that is usually conducted in order to retrieve lexical bundles (see Biber et al. 1999). A 2-gram would consist of two words, a 3-gram of three words and so on. 3-grams are extremely frequent in spoken discourse as displayed in figure 40 (Biber et al. 2002: 444).

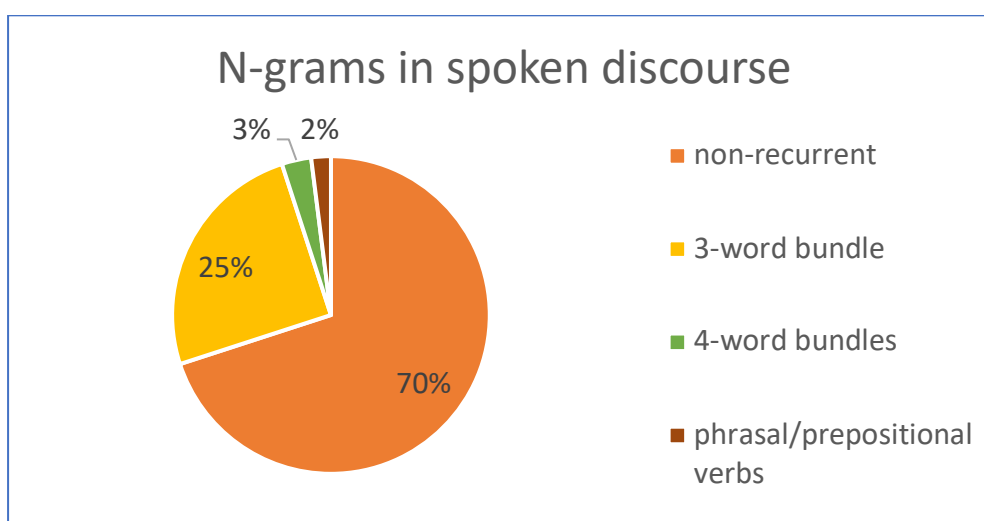


FIGURE 40: FREQUENT N-GRAMS IN SPOKEN DISCOURSE (ADAPTED FROM BIBER ET AL. 2002: 444)

Biber and colleagues (2002) found that most of the word combinations in spoken discourse are non-recurrent expressions (Biber et al. 2002: 446). One quarter of the expressions in spoken discourse are 3-grams, followed by 4-grams, which make up only 3% of the expressions in spoken discourse. Since 3-grams and 4-grams are very common in spoken language and make up of a large part of spoken discourse (see figure 40), they will be investigated further in the following.

In terms of patterns of lexical bundles, Biber et al. (2002: 446) came up with three pattern categories for lexical bundles in conversations:

TABLE 25: COMMON LEXICAL BUNDLES IN CONVERSATION

Personal pronoun + verb phrase + XY	<i>I don't know if</i> <i>I would like to</i> <i>I tell you what</i>
Extended verb phrase fragments	<i>Have a look at</i> <i>Know what I mean</i>
Question fragments	<i>Do you want to</i> <i>Do you know what</i>

In EFL teaching, these multi-word sequences or lexical bundles are usually called (lexical) chunks, which usually includes lexical bundles, collocations and idioms. Lexical (chunks) are very interesting for EFL teaching and found to be extremely useful for learners (see for example Biber et al. 1999; Hunston & Francis 2000; Lewis 2002; and Sinclair & Renouf 1988). How these lexical bundles can be used in EFL teaching and learning will be further explored in section 5.1.1. In terms of selection of lexical bundles, frequency provides a good measure, which is usually investigated in a corpus-based approach. In including frequency data and analysis, “patterns of use that often go unnoticed by researchers” are made visible and “must be explained” (Biber et al. 2004: 376). Therefore, frequently occurring lexical bundles should be investigated and are worth teaching (see section 5.1.1.).

Lexical bundles have been analyzed in various studies with regard to their frequency, but also with regard to their functions (see for example Biber et al. 2004, Cortes 2004, Hyland 2008 and Simpson-Vlach & Ellis 2010). Biber et al. (2004) investigated lexical bundles in classroom teaching and textbooks

and compared it to the use of lexical bundles in conversation and academic prose, thus they looked at the use of lexical bundles across different registers – in speech and writing. Based on the frequency list, they established a taxonomy for different functions of lexical bundles (Biber et al. 2004: 389 ff):

- **Stance bundles:** epistemic stance bundles and attitudinal / modality stance bundles
- **Discourse organizing bundles:** topic introduction/focus bundles; topic elaboration/clarification bundles
- **Referential bundles:** identification/focus, imprecision indicators, specification of attributes, time/place/text reference) Stance bundles

Simpson-Vlach and Ellis (2010) collected an “empirically derived, pedagogically useful list of formulaic sequences for academic speech and writing”, the **Academic Formulas List (AFL)**. They used Biber et al.’s (2004) taxonomy and adapted it so that the main categories remain the same, but the sub-categories are specified. For example, Simpson-Vlach and Ellis (2010) identified new functions such as hedges and expressions of ability and possibility for the category of stance expressions.

The AFL function categories are the following (Simpson-Vlach & Ellis 2010: 498ff):

- **Group A: Referential expressions:** Specification of attributes, identification and focus, contrast and comparison, deictics and locatives
- **Group B: Stance expressions:** Hedges, epistemic stance, obligation and directive, expressions of ability and possibility, evaluation
- **Group C: Discourse organizing functions:** Meta-discourse and textual reference, topic introduction and focus, topic elaboration and discourse markers

Another widely used taxonomy for lexical bundles was established by Hyland (2008), who also based his functions on Biber et al. (2004). He looked at lexical bundles in academic written texts (research articles, doctoral dissertations, Master’s theses) and investigated frequently used 3-grams, 4-grams and 5-grams. In the discussion of the results, he differentiated between research oriented-function, text-oriented function and participant-oriented function. This categorization differs in terms of main categories from Biber et al. (2004) and Simpson-Vlach & Ellis (2010). As far as research-oriented function is concerned, he lists different aspects such as location, procedure, quantification, description and topic. Among the text-oriented functions, he states that lexical bundles are used as transition

signals, resultative signals, structuring signals or framing signals. Participant-oriented lexical bundles are used to express stance features or engagement features (Hyland 2008: 13ff). There are some functions, such as stance features, that can be found in all three categorizations. Another aspect that is similar in all studies is that they started with a frequency analysis of 3-grams and 4-grams and then classified them according to functions. This methodology is used in this study as well. As far as the functions are concerned, Biber et al. (2004) and Simpson-Vlach and Ellis (2010) taxonomy will be used, as they investigated lexical bundles spoken corpora (besides written corpora). Their results can be better compared to the lexical bundles in TaCoCASE, which is also a spoken corpus.

Having set the theoretical background, I will discuss the findings of the collocation and n-gram analysis conducted with TaCoCASE in the following.

Collocation Analysis

Really is the most frequent adverb in TaCoCASE – rank 32 according to the frequency word list created with WebCorpLSE – and it appeared in each of the 15 conversations. The overall frequency of *really* is 2,150 pmw in TaCoCASE. *Really* is a highly frequent adverb in spoken discourse that has been investigated by many researchers (see for example Stenström 1986, Paradis 2003, Biber and Staples 2014, Pérez-Paredes & Bueno-Alastuey 2019, Diani 2008). In general, adverbs seem to be very difficult to acquire for learners as they have different functions (see for example Götz & Schilk 2011; Lorenz 1999).

Really can be used as an epistemic stance adverbial, expressing a degree of truth and factfulness, usually in clause final or initial position (Biber et al. 2002: 385; see also Paradis 2003 and Pérez-Paredes & Bueno-Alastuey 2019). It can also be used as an intensifier, usually in combination with verbs or adjectives and usually in clause medial position (Biber et al. 2002: 385). Additional functions that have been discovered, mainly by Pérez-Paredes and Bueno-Alastuey (2019) in research among NS and NNS are: “booster (intensifying function), skeptical response to something said earlier, doubt or surprise as

well as factually true meanings” (see also Myers 2010). Pérez-Paredes and Bueno-Alastuey (2019) found that *really* is significantly more frequently used by German speakers compared to English native speakers.

WebCorplSE was used to examine collocates of *really* in TaCoCASE. A collocate search was conducted with a word span of 4 words to the left and the right of the keyword. The results were ordered by log-likelihood score (column 7), which provides in this case information about the strength of the collocate, that was measured with WebCorplSE. The higher the score, the stronger the co-occurrence and thus the less likely it is that the two words appeared together by chance.

TABLE 26: COLLOCATES OF *REALLY* IN TACOCASE

Rank	Lemma	POS	Left Freq.	Right Freq.	Total	Score
1	REALLY	ADV	41	41	82	363.133
2	NICE	ADJ	2	33	35	216.540
3	BIG	ADJ	4	19	23	123.733
4	OH	INTJ	29	5	34	93.233
5	COOL	ADJ	2	16	18	90.815
6	JUST	ADV	18	18	36	83.161
7	UH	INTJ	22	13	35	74.642
8	UHM	INTJ	8	13	21	65.943
9	INTERESTING	ADJ	3	12	15	65.595
10	BAD	ADJ	1	12	13	62.879
11	UHM	INTJ	13	20	33	62.849
12	KNOW	VERB	8	23	31	59.361
13	GOOD	ADJ	3	17	20	59.246
14	GREAT	ADJ	3	8	11	58.351
15	THINK	VERB	16	10	26	50.849
16	HARD	ADJ	2	8	10	48.430
17	DEPEND	VERB	3	7	10	48.093
18	EXCITED	ADJ	2	4	6	46.425
19	ENJOY	VERB	0	7	7	40.546
20	ACTUALLY	ADV	11	5	16	40.005

As the results in table 26 indicate, most of the collocates of *really* are adjectives preceded by *really* such as *nice*, *big*, *cool*, *interesting*, *bad*, *good*, *great*, *hard*, *excited*. It is interesting to see that in most cases, the adjectives are used on the right of *really*, for example *really nice*, *really big*, *really cool*. This

means that *really* is mostly used in the syntactic structure “*really* + ADJ”, which draws on its function as modifier, intensifier or booster (see Biber et al. 2002; Pérez-Paredes & Bueno-Alastuey 2019). Another frequently used syntactic structure is “*really* + VERB”, which also draws on its function as modifier (see Biber et al. 2002; Pérez-Paredes & Bueno-Alastuey 2019). The verb *enjoy* only appears on the right of *really* and not at all on the left, which means that the combination *really enjoy* is far more frequent than the combination **enjoy really*. The same tendency can be found with the verb *know*, where *really* is used in 2/3 of the cases in front of the verb *know*: *really (...) know*, or as expressed in the following example: *I don’t really know much about the Scottish accent*. Interestingly, *really* collocates with itself 82 times, which makes it a very frequently used adverb in TaCoCASE. This repeated use of *really* functions as emphasis or intensifier, as the following example shows: “*a family friend of ours uhm was really suffering from really really bad depression*” or “*on Thursday, I’ve got two seminars for a really really big exam*” (see Biber et al. 2002; Pérez-Paredes & Bueno-Alastuey 2019).

Thus, *really* is used in combination with adjectives and verbs, usually as modifier and usually in front of adjectives and adverbs. It is used to express emphasis or truthfulness and functions as modifier, intensifier or booster in combination with adjectives and verbs (see Biber et al. 2002; Pérez-Paredes & Bueno-Alastuey 2019).

N-gram analysis (lexical bundles)

I follow Biber et al. (2004) in using the term **lexical bundle or n-gram**, as these terms are usually used in corpus linguistics to extract or investigate multi-word units. As I mainly use corpus methods to analyze the multi-word sequences, I chose to use these terms that are common in this research area. In this case, the n-grams were extracted with AntConc and the n-grams under investigation were 3-grams and 4-grams as they are very frequent in spoken discourse (Biber et al. 2002: 444). The range was set to 15 for the 3-grams as the n-grams should appear in every conversation in TaCoCASE. The range for the 4-grams was set to 10, because there were only two reasonable results with a range of 15, and therefore the range was lowered to 10 conversations.

TABLE 27: 3-GRAMS IN TACOCASE

Type	Rank	Freq	Range
i don t	1	532	15
don t know	2	316	15
a lot of	3	204	15
it s like	4	134	15
it s a	5	110	15
yeah it s	6	105	15
i m not	7	104	15
it s just	8	97	15
i think it	9	92	15
it s not	10	91	15

TABLE 28: 4-GRAMS IN TACOCASE

Type	Rank	Freq	Range
i don t know	1	292	15
it s it s	2	80	12
don t know if	3	68	13
i think it s	4	66	15
like i don t	5	38	10
it s kind of	6	36	11
a lot of people	7	34	12
i don t have	7	34	13
i don t think	9	31	13
i m i m	9	31	13

Table 27 displays the top 10 3-grams in TaCoCASE and table 28 displays the top 10 4-grams in TaCoCASE. The most frequent 3-gram is *I don't* (532) and the most frequent 4-gram is *I don't know* (292). This is a pattern “personal pronoun + verb phrase + XY” that Biber and colleagues (2002: 446) also found to be very frequent in conversations (see table 25). Besides *I don't know*, the pattern *I don't* plus a verb like *know*, *have* or *think* is also quite frequently used in TaCoCASE.

Other frequent expressions is *it's* XY, such as *it's like*, *it's a*, *it's just*, *it's kind of*, *it's no*, which also falls into the pattern category “personal pronoun (*it*) + verb phrase (*is*) + XY (*like*, *just*, *kind of*)”. The other two patterns that Biber and colleagues found “extended verb phrase” and “question fragments” (see

table 25) are not in the top 10 3 or 4-gram list in TaCoCASE. However, the quantifier *a lot of* (204; see table 27) is also frequently used and usually followed by a noun, in this case *a lot of people* (34; see table 28).

It is interesting to see that 3-grams and 4-grams are very similar and that 4-grams are usually extended 3-grams. This supports findings by Biber et al. (1999: 992), who found that larger lexical bundles are usually related to smaller bundles.

In terms of **functions**, most of the lexical bundles found in TaCoCASE are stance bundles (Biber 2004, Simpson-Vlach & Ellis 2010). In fact, most of the stance bundles can be classified as personal epistemic bundles, such as *I don't know if/what/how* and *I (don't) think* (see Biber et al. 2004). *I don't know if/what/how* is a personal epistemic stance bundle used to express uncertainty (see also Biber et al. (2004: 389). *I (don't) think* is used to express “possibility, but lack of certainty” (Biber et al. 2004: 389). Another stance bundle is *it's kind of* (rank 6; table 27). It is used as a hedge to express ambiguity and probability (see also AFL, Simpson-Vlach & Ellis 2010).

The lexical bundle *a lot of* is the only referential expression found in TaCoCASE. It is used to specifying quantity (see Biber et al. 2004: 395, see also Simpson-Vlach & Ellis 2010).

It is interesting to see that the lexical bundles in TaCoCASE represent two of the main functions identified in previous research – **stance expressions and referential expressions** (see Biber et al. 2004; Simpson-Vlach & Ellis 2010). The third category, discourse organizing expressions (see Biber et al. 2004; Simpson-Vlach & Ellis 2010) was not found among the top 10 lists of 3-grams and 4-grams in TaCoCASE (in fact, also not in the top 50 list). This is quite surprising as discourse organizing expressions are usually very frequently used in conversation (see Biber et al. 2004). It seems that the participants use other strategies to organize discourse than using lexical bundles, for example discourse markers (see section 4.1.4.). It could be the case that due to the multimodal setup and the computer-mediated communication, the participants do not need lexical bundles to organize discourse and that they use multimodal features, such as gestures. In fact, gestures and non-verbal elements, such as *head to the*

right (rank 18) or *looks to side* (rank 20) are quite frequent bundles in the top 50 list of 4-grams in TaCoCASE, even though they do not have discourse organizing functions. Using discourse markers is another possibility to structure and organize discourse, which the participants made use of (see section 4.1.4.). It seems that the participants favored using single words as discourse markers or discourse marker combinations, which explains why we do not find lexical bundles as discourse markers. The use and functions of two single word discourse markers, that have discourse organizing functions – *so* and *okay* (see for example Johnson 2002; Buysse 2012; House 2013) – will be investigated in the following section (see section 4.1.4.).

The most frequent expression *I don't know XY* will be analyzed in greater detail in the following. It is not very surprising that *I don't know* is a frequent n-gram as it was found to be the “single most frequent negative collocation” in COCA and the BNC (Baumgarten & House 2010: 1186). It is usually used as an answer if the interlocutor is not able to provide the information that he or she is being asked about (Tsui 1991). Functions of *I don't know* that have been stated in previous research include: avoidance of making an assessment, preface to a disagreement, an avoidance of an explicit disagreement, avoidance of commitment, minimization of impolite beliefs and marker of uncertainty (see Tsui 1991; Baumgarten & House 2010). Despite these different functions, Tsui (1991) states that *I don't know* has a “unifying meaning”, that is “insufficient knowledge” (Tsui 1991). *I don't know* is usually used by native speakers as well as non-native speakers to “construct a stance” (Baumgarten & House 2010: 1198). Native speakers are found to use *I don't know* to express “uncertainty, avoidance, neutrality and non-commitment”, whereas non-native speakers usually use *I don't know* as a “marker of insufficient knowledge” and “use the expression to verbalize and to overcome on-line planning difficulties”. (Baumgarten & House 2010: 1198). Thus, it also has a pragmatic motivation, namely face-saving – either of the person itself or the other.

A wildcard search in the n-gram analysis was conducted using the node *I don't know **, where the asterisk symbolizes the wildcard. Table 27 displays frequently used expressions with *I don't know XY*

in the concordances. The concordances are ranked in alphabetical order and not by frequency so that patterns become visible.

TABLE 29: CONCORDANCES OF *I DON'T KNOW*

No.	Conv. ID	Concordance		
1	01SB151BO05	I ever e:r, do desire to go to, Europe,	I don't know if	I actually would want to, ...
2	07SB150BI32	called Muttertag ((German (1.0))), so Mother Day to translate it literally,	I don't know if	you have that as well?
3	08BO15BI36	BI36: [okay]. BO15: [we've], .. bu:t zumba, I-	I don't know if	you know what zumba is?
4	08BO18BI41	BO18: [mh]. {nods} BI41: (1.5) Treasure Island?	I don't know if	you heard of Treasure Island.
5	02BI33BO10	BO10: ((eh)) yeah ((/ya/)) I'm not. {shakes head}	I don't know what	it is. {raises eyebrows} but .
6	03SB149BO03	train, for like ten minutes before it would leave. ((chuckles))	I don't know what	was wrong with it but. ...
7	01SB151BO05	BO05: .. in Asia, ... uhm, (1.0)	I don't know how	the pay compares to,
8	05SB124BI14	BI14: mhm. SB124: and actually to .. to uhm, (2.3) like	I don't know how	to say it right now,
9	06SB157BO11	places here, that make like authentic style ramen ((Japanese (0.5))), but	I don't know how	authentic it is, {shakes head}
10	07SB150BI32	I wouldn't agree on that, to be honest and	I don't know how	it's in, England but

The examples in table 30 follow the same syntactic structure:

first person singular pronoun *I* + negative auxiliary verb (*don't* as contraction of *do not*) + main verb (*know*) + complement.

As the examples in table 28 suggest, *I don't know* is used as a stance-marker (see Baumgarten & House 2010, Tsui 1991). The complement introduces the aspect the speaker is uncertain about. The

complement can be introduced by **interrogative pronouns** such as *how* and *what* was found in examples (5) *I don't know what it is*, (6) *I don't know what was wrong with it*, (7) *I don't know how the pay compares to...*, (8) *I don't know how to say it right now*, (9) *I don't know how authentic it is*, and (10) *I don't know how it is in England*. It could also be introduced by the **conjunction if**, introducing an if-clause or a conditional clause as found in examples (1) *I don't know if I actually would want to...*, (2) *I don't know if you have that as well*, (3) *I don't know if you know what Zumba is* or (4) *I don't know if you heard of Treasure Island*.

The **main functions** that were identified in previous research – **expressing uncertainty and insufficient knowledge** (see Tsui 1991; Baumgarten and House 2010) are the functions found in TaCoCASE and the concordances in table 28, for example (6) *I don't know what was wrong with it* or (7) *I don't know how the pay compares to*. It is interesting to see that the uncertainty in some cases results from lack of knowledge of cultural aspects or concepts (see examples 2, 3, 4, 7, 9, 10). In examples 2, 3 and 4 the speakers ask in an indirect but polite way, if the interlocutor knows a specific concept or if they have seen a specific TV series. The speaker does not know whether the interlocutor knows the concept he or she is introducing and thus signals insufficient knowledge. It is interesting that it is used as part of a question, whereas *I don't know* has previously been found to be an answer to a question (see Tsui 1991). This question is motivated by a face-saving act towards the interlocutor (see Tsui 1991) when a speaker is not sure if the interlocutor knows a cultural concept or not. The speaker implies that he or she assumes that the interlocutor does not have to know the concept and that he or she is aware that cultural concepts cannot be taken for granted. In a similar manner, examples 7, 9 and 10, are also motivated by face-saving acts, as they are concerned with authenticity of food (ramen), norms in England and the pay in Japan. All these examples symbolize a lack of knowledge or uncertainty of the respective culture they are talking about. In this case, the face-saving act is towards a culture and not the interlocutor itself. This is quite interesting as this shows cultural sensibility and cultural awareness, two factors that are important in intercultural competence (see Byram 1997, 2021; see also Barmeyer 2000). This is a new function that has not appeared in previous research – to the best of my knowledge

– and that would be worth investigating further in future research. Section 4.2.1. will have a closer look at what strategies the participants used to explain (new) concepts. It is clear from this analysis that the most frequently used expression – from a linguistic perspective – is connected to intercultural aspects and concepts on the content level, which underpins the fact that language and culture are interconnected.

4.1.4. DISCOURSE MARKERS

RQ4: What are frequently used discourse markers in TaCoCASE and how are they used?¹⁹

Discourse Markers (DM), sometimes also referred to as linking devices or (linking) connectors, are “sequentially dependent elements which bracket units of talk, e.g. *oh, well, and, but, or, so, because, now, then, I mean, y’ know*” (Schiffrin 1987: 31). They are used to “signal a sequential relationship between the current basic message and the previous discourse” (Fraser 1990: 383). Even though they are “grammatically optional and semantically empty” (Brinton 1996), they play an important role in terms of pragmatics, “the study of ‘invisible’ meaning” (Yule 2006: 112). Discourse markers contribute to cohesion – the underlying structure and connectedness of a text (Yule 2006: 125) and coherence, meaning “everything fitting together well” in a text (Yule 2006: 126). Discourse markers can have different functions when used in native speaker conversations in comparison to non-native speaker or ELF conversations (Buysse 2012; House 2013, Collet et al. 2021). When examining discourse markers, it is important to look at DM within action sequences (Bolden 2009, Aijmer 2002). The DM function is usually tied to the position within the sequence and also the position of the sequence within larger interaction (Bolden 2009, Raymond 2004). As Koops and Lohmann (2016) found, DM frequently co-occur with other DM and interact as part of a sequence or two or more DM. According to them, DM

¹⁹ This section draws on my previous article that investigated discourse markers in ViMELF: Collet, C. et al. (2021). So in video-mediated communication in the Expanding Circle. *World Englishes* 40(4), 1-17.

combinations fulfill different functions than the single DM. DM also co-occur with paralinguistic and non-verbal elements (see Goodwin & Goodwin 2000). Research has shown that discourse markers pose difficulties for learners and that they are difficult to acquire, which sometimes results in problematic uses of discourse markers, such as overuse or misuse (see Milton & Tsang 1993, Hellerman & Vrgun 2007, Leńko-Szymańśka 2008, Chen 2006a, Collet & Ankerstein 2015). In order to establish communicative competence, it is important to build pragmatic competence and discourse competence, which includes teaching about discourse markers, their use and their functions.

One very versatile discourse marker that fulfills many functions is **so**. Research on *so* as discourse marker has been conducted among native and non-native speakers and **different functions of *so*** were derived.

- On the **ideational level**, *so* is used to mark inference (Blakemore 1988, Fraser 1999) or result (Schiffrin 1987, Buysse 2012). In NNS talk, *so* is used self-attentive and displays the speaker's mental process (Fischer 2006).
- On the **interpersonal level**, *so* functions as marker of emergence from incipency, and is used in NS conversations always other attentive (Bolden 2006). It is used as a prompt for the interlocutor to produce the next action (Raymond 2004, Buysse 2012). In NNS talk, *so* is found to be used as a hesitation or transition marker (Redeker 2006, Lam 2010) or as holding the floor device (Buysse 2012).
- On the **textual level**, *so* is used as a topic developer or topic sequencer (Johnson 2002, Buysse 2012) and as marker of connection (Howe 1991). In an ELF context, House (2013) found that *so* is used speaker-supportive as a marker of (inter)subjectivity or connectivity. Redeker (2006) identified deictic uses of *so*, meaning that it can point forward to indicate new moves and backwards to sum up what had just been said. Buysse (2012) found that *so* is used in terms of self-correction and for elaboration.

These documented functions of *so* in NS and NNS or ELF contexts are summarized in table 30 (see Collet et al. 2021).

TABLE 30: DOCUMENTED FUNCTIONS OF *SO* IN NS AND NNS CONTEXTS

Ideational	Interpersonal	Textual
Result; inference; self-attentive (displays speaker's mental process)	Marker of emergence from incipency; prompt for interlocutor to produce next action; hesitation or transition marker; holding the floor device	Topic developer and topic sequencer; marker of connection; deictic function (new moves and summing up); self-correction; elaboration

Comparing the documented functions of *so* among native and non-native speakers, it can be stated that *so* fulfills the same basic functions, such as result, prompt, topic developer, in NS and NNS conversations. Some additional functions of *so* were identified among NNS, which are often self-attentive in combination with hesitation or self-correction. In NNS conversations, there could also be an L1 influence, for example of the German *so*, which means that the L1 functions of *so* could be transferred in their English conversation.

The collocate analysis of *so* revealed a strong co-occurrences of *so* as DM with *okay* as DM. Therefore, we will also take a look at *okay* when it functions as DM in TaCoCASE and the combination of both. Documented functions of *okay* as DM in NS contexts include pre-close (Schegloff & Sacks 1973) and closing marker (Edmondson & House 1981, Fung & Carter 2007), as well as backchannel (Swales & Malczewski 2001, Fung & Carter 2007) and marker of return to topic after digression (Bangerter & Clark 2003). Fung & Carter (2007) found that *okay* as DM is used in agreeing, to signal topic change and as comprehension check. A greater variety of functions had been identified in ELF contexts by Juliane House (2013: 64). She found that *okay* as used as DM functions as opener or starter and attention getter, as linking device, general check, signal of approval (consent), pre-close and closing signal as well as sealing outcome signal (Edmondson & House 1981).

The **use and functions** of *so* as a discourse marker in TaCoCASE will be examined taking aspects such as frequency and collocates into account. It will be examined to what extent the use of DM is different among NS and NNS and what functions the DM fulfill. In a last step, DM combinations will be investigated.

So fulfills many functions in discourse such as adverb, conjunction or intensifier depending on its role within a sentence (see Biber et al. 2002). Since only DM functions are of interest in this work, the corpus was manually tagged, using <D=*so*> to indicate that *so* was used in DM function and <X=*so*> for all other uses. After that, a quantitative frequency analysis and a collocation analysis was carried out. The different functions were identified. Following a corpus-based discourse analytical approach with elements from conversation analysis, a qualitative analysis was carried out in order to investigate what functions *so* fulfills in discourse and how the speakers used *so* in their conversations with regard to discourse structure and turn-taking behavior. In order to have a larger database and for reasons of comparison, data from ViMELF was used where necessary.

Frequency Analysis²⁰

So as a DM appears in all 15 conversations with a raw frequency of 1,765 and a normalized mean frequency of 11.09 per 1,000 words across the different speakers. The most frequent DM that co-occurs with *so* is *okay* (LL = 43.365). Therefore, we will take a look at *so* and *okay* as DM in the following.

Table 31 displays the overall **frequency** of *so* and *okay* when used as discourse marker or with other functions in TaCoCASE and table 32 in ViMELF.

²⁰ Note that for the quantitative analysis, a preliminary version of TaCoCASE had been used, which differs slightly in total number of words. This means that the results for the normalized frequency are slightly different than when using the total number of words of the published TaCoCASE version.

TABLE 31: FREQUENCY OF *SO* AND *OKAY* IN TACOCASE

TaCoCASE	<D=so>	<X=so>	<D=okay>	<X=okay>
raw frequency	1,393	372	532	27
mean frequency per 1,000 words	11.09	2.73	5.13	0.21
percentage per category	78.92%	21.08%	95.17%	4.83%

TABLE 32: FREQUENCY OF *SO* AND *OKAY* IN ViMELF

ViMELF	<D=so>	<X=so>	<D=okay>	<X=okay>
raw frequency	1,417	271	822	45
mean frequency per 1,000 words	10.87	2.01	6.22	0.34
percentage per category	83.95%	16.05%	94.81%	5.19%

So appears as DM 1,393 times in TaCoCASE (mean normed frequency = 11.09 per 1,000 words) and 372 times with other functions, which means that *so* is used in DM functions in 78.92 % of the cases. *So* appears as DM in ViMELF 1,417 times (mean normed frequency = 10.87 per 1,000 words) and 271 times with other functions, which means that in 83.95% of the cases, *so* is used as a DM. *Okay* is used about 5.13 times per 1,000 words and in 95.17% of the cases as a discourse marker in TaCoCASE and 6.22 times per 1,000 words and in 94.81% of the cases as discourse marker in ViMELF.

It is interesting to see that the overall mean frequency of *so* when used as a discourse marker is lower in ViMELF (10.87 per 1,000 words) compared to TaCoCASE (11.09 per 1,000 words). This means that *so* in its function as discourse marker is used more frequently in TaCoCASE than in ViMELF, while *okay* is used more frequently as discourse marker in ViMELF (6.22 per 1,000 words) than in TaCoCASE (5.13 per 1,000 words).

The following two figures show the **distribution per category** of *so* and *okay* for reasons of comparison.

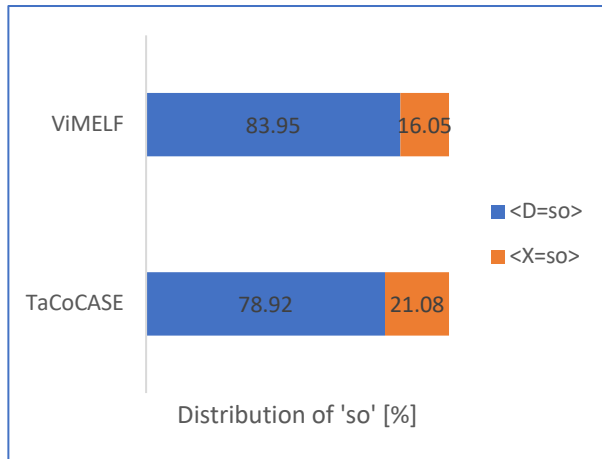


FIGURE 41: DISTRIBUTION OF *so* PER CATEGORY

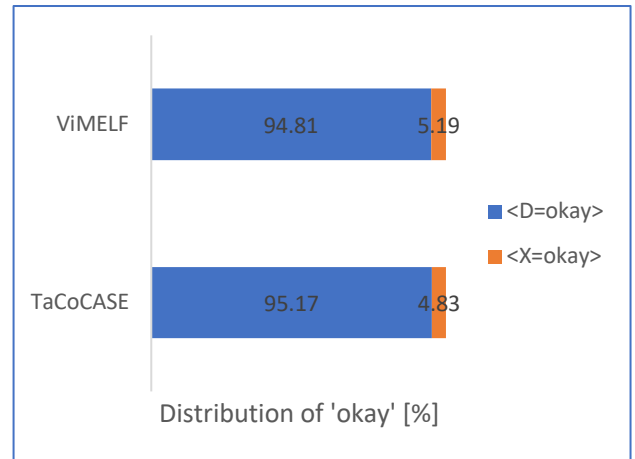


FIGURE 42: DISTRIBUTION OF *okay* PER CATEGORY

While *okay* is used with quite the same percentage per category as discourse marker in ViMELF (94.81 %) as well as in TaCoCASE (95.17%), the numbers are different for *so* as discourse marker. While *so* is used in 78.92 % as DM in TaCoCASE, it is even more frequently used as DM in ViMELF (83.95%) compared to its non-DM uses. This means that the distribution of function as discourse marker is quite stable for *okay* as discourse marker in both corpora, but differs slightly for *so*.

It is worth taking a closer look at TaCoCASE and the speaker groups (BI, SB, BO) in order to compare the use of DM between NS and NNS.

Figure 43 displays the mean frequency of *so* as DM for the British speakers (BI), the German speakers (SB) and the American speakers (BO).

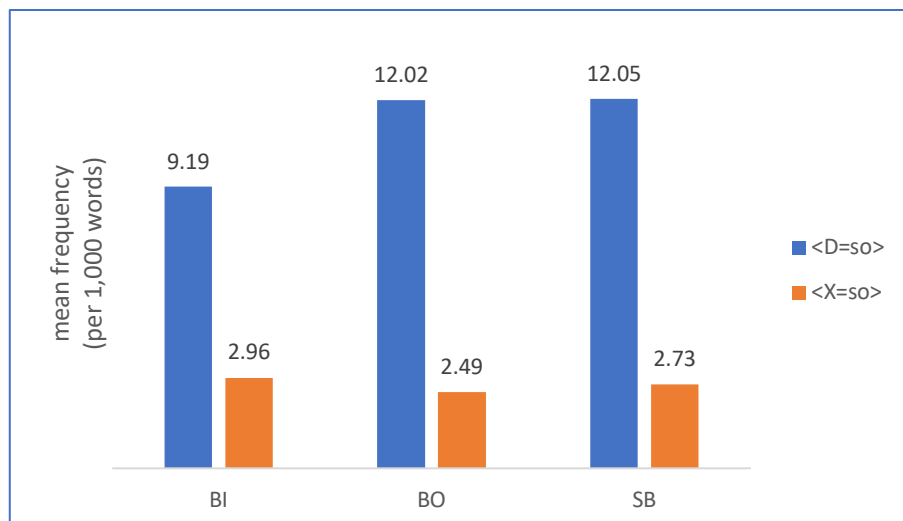


FIGURE 43: FREQUENCY OF *so* ACROSS NS AND NNS IN TACOCASE

Interestingly, the numbers for the American and German speakers are quite close as the American speakers used *so* as DM 12.02 times per 1,000 words and the German speakers used *so* as DM 12.05 times per 1,000 words. There is a bigger difference between the American as well as the German and the British speakers, as the British speakers used *so* as DM only 9.19 times per 1,000 words. From looking at the data, it is clear that calculating the mean number for BI and BO speakers and looking at NS versus NNS is not very useful as the deviation is too high. Therefore, a general statement to compare NS and NNS use is not possible, even though it can be stated that the German speakers and the American speakers used *so* in a similar way and with similar frequency, whereas the use differs between the British speakers and the German speakers.

Figure 44 displays the mean frequency of *okay* when used as DM for BI, BO and SB speakers in TaCoCASE.

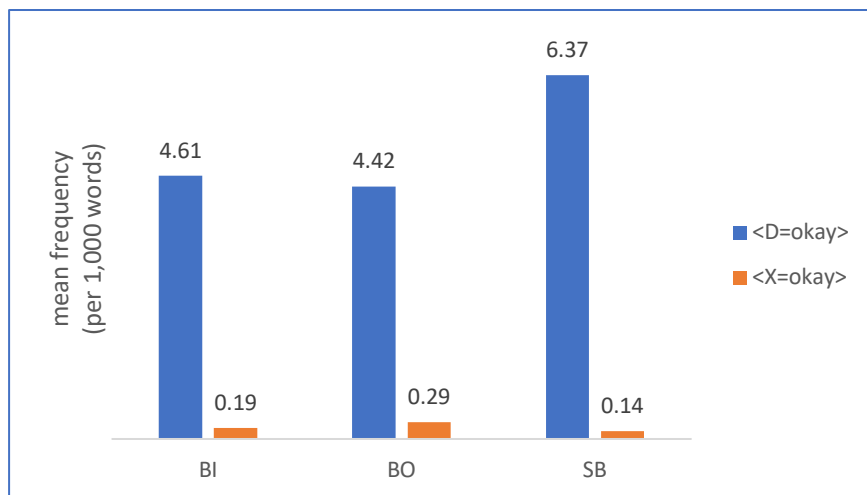


FIGURE 44: FREQUENCY OF *OKAY* ACROSS NS AND NNS IN TACOCASE

The use of *okay* as DM differs not so much between BO and BI speakers as the BI speakers used *okay* as DM 4.61 times per 1,000 words and the BO speakers used it 4.42 times per 1,000 words. The number for the German speakers is much higher as they used *okay* as DM 6.37 times per 1,000 words. In this case, a difference between the NS use and the NNS of *okay* as DM in TaCoCASE use can be stated.

However, mean frequencies across speaker groups are not too meaningful and looking at the individual speakers and their use could unveil new findings. Figures 45 and 46 display the mean frequency of *so* or *okay* as DM per individual speaker in TaCoCASE. The bars in figures 45 and 56 are colored orange (BI), green (SB) and blue (BO) so that the different speaker groups have the same color. The use of *so* as DM varies from the least frequent use (1.06 per 1,000 words) by BI40 to the most frequent use (19.68 per 1,000 words) by BO06. *Okay* as DM is used with the lowest frequency by BI29 (0 per 1,000 words), followed by BI14 (0.17 per 1,000 words) to the most frequent use by SB147 (23.48 per 1,000 words). Interestingly, the use of DM is even different for the same speaker but in a different conversation. SB156 used *so* as DM in the conversation with BO07 (4.62 per 1,000 words) less frequently than in the conversation with BI03 (7,18 per 1,000 words) and there are also differences for BO18, BO10 and SB 149, who held two conversations with two different conversation partners.

When looking at the individual speakers in TaCoCASE, it becomes clear that the use of the DM is very different and varies for every individual speaker. A trend for speaker groups (BI, BO, SB) cannot really be identified.

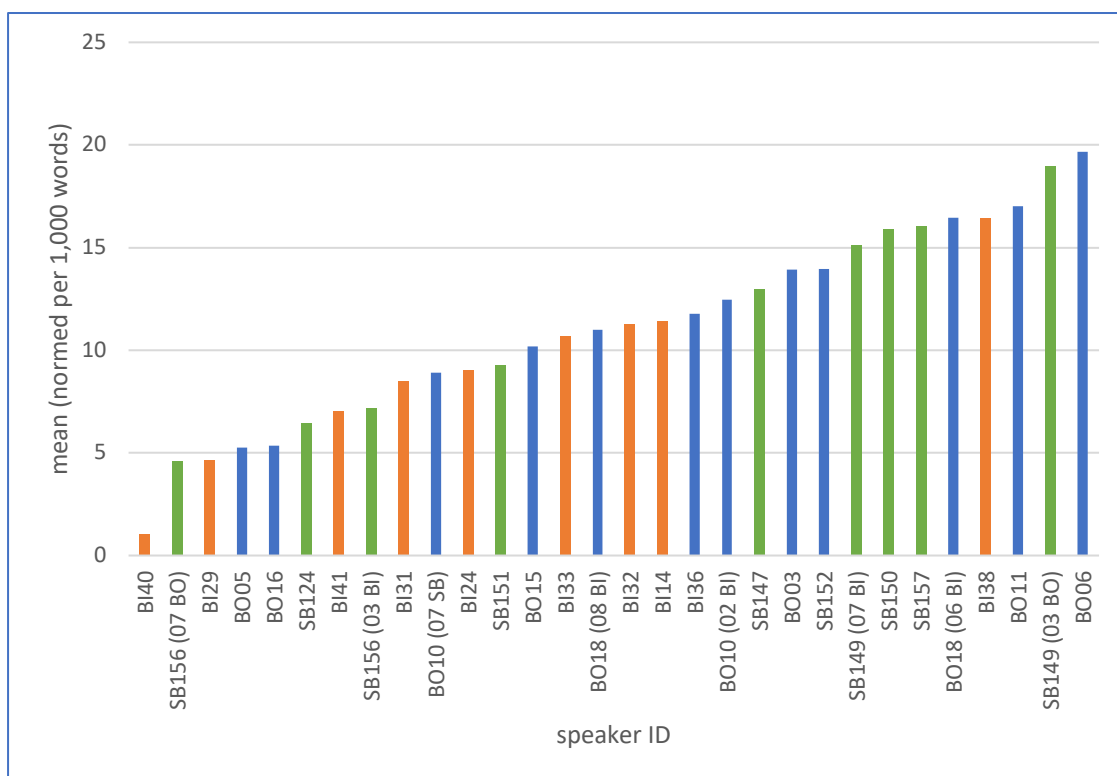


FIGURE 45: DISTRIBUTION OF *SO AS DM* ACROSS INDIVIDUAL SPEAKERS IN TACOCASE

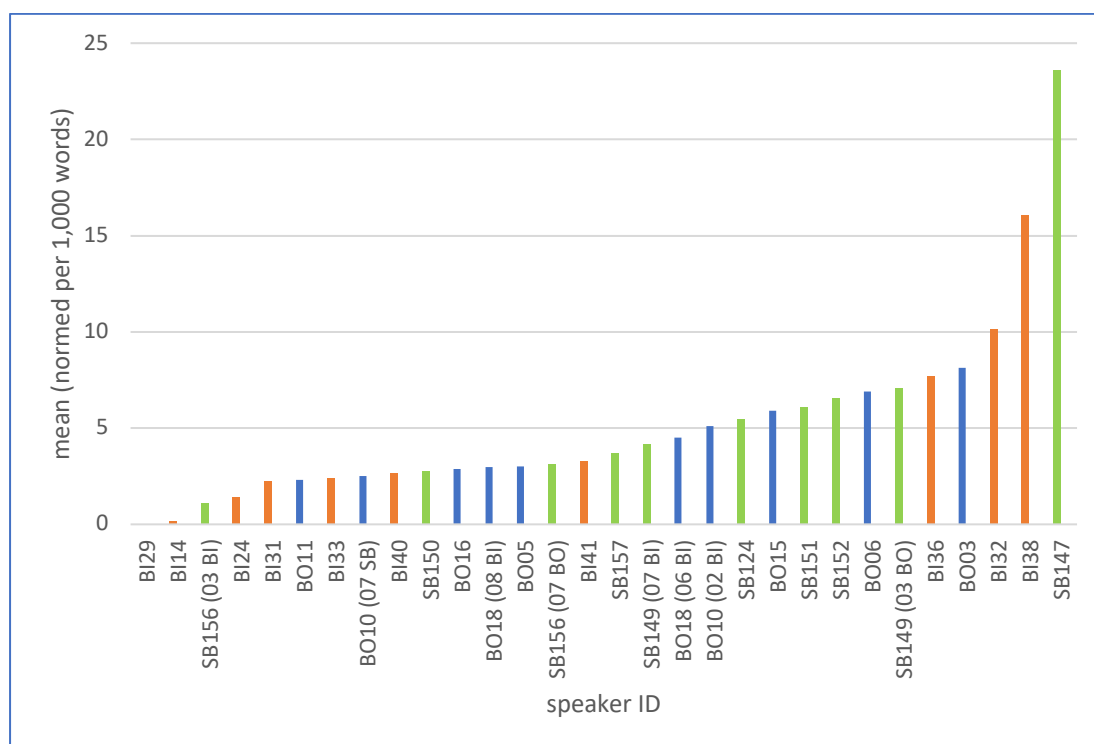


FIGURE 46: DISTRIBUTION OF *OKAY AS DM* ACROSS INDIVIDUAL SPEAKERS IN TACOCASE

Collocates of *so*

I conducted a **collocate search for *so***, using AntConc, and setting the word span to 2L and 2R. The 15 strongest collocates of *so* are displayed in table 33. The results are ordered by likelihood as this is a measure of the strength of the co-occurrence.

TABLE 33: COLLOCATES OF *SO* IN TACOCASE

Collocate	Rank	FreqL	FreqR	Range	Likelihood
many	1	2	45	12	148.523
s	2	112	155	15	88.584
much	3	4	40	13	73.095
to	4	55	19	14	59.706
a	5	22	28	14	55.769
i	6	70	337	15	53.605
okay	7	50	17	11	43.365
uhm	8	59	49	14	40.049
of	9	27	16	12	39.422
it	10	111	135	15	38.069
sb	11	156	74	10	35.681
the	12	38	46	13	35.054
different	13	10	18	10	29.053
in	14	18	15	12	26.291
bo	15	130	85	10	25.590

Many (rank 1) and *much* (rank 3) are strong collocates of *so*, but *so* in this case is used as an intensifier and not as DM. The same is true for the collocate *different*, where *so* is also used as an intensifier. The strongest DM collocate of *so* is *okay* (rank 7), which is the same for ViMELF (see Collet et al. 2021: 601). In terms of sequence preference, *okay* is used more frequently on the left of *so* – “*okay .. so*” – and less frequently on the right “*so .. okay*”. Interestingly, the speaker IDs *SB* and *BO* are also among the top 15 strongest collocates. This indicates that *so* is frequently used turn initial, if the speaker ID precedes *so*, or turn final, if the speaker ID follows *so*.

Functions of *so* as DM in TaCoCASE

A previous study (Collet et al. 2021) demonstrated that all seven main functions of *so* documented in previous research – result and inference, action prompt, topic developer and sequencer, connection marker, marker of emergence from incipency – could be identified in ViMELF (Collet et al. 2021: 602 f). The quantitative results suggest that the use of *so* in the two corpora is quite similar and that the most frequent DM as collocate of *so* is *okay* (see above). A similar qualitative analysis was conducted with TaCoCASE and the same methodology as the previous study with ViMELF was applied (see Collet et al. 2021: 598). The DM-tagged version of TaCoCASE was used for the qualitative analysis and the main documented functions were interpreted from the context in the conversations. This means that I looked at the close context as well as the wider discourse including conversational structure, discourse strategy and the general conversational setting in order to identify and label the different functions. A similar approach was chosen for the DM combination of *okay so*, that will be described in the following.

The following functions of *so* were found in TaCoCASE:

- (1) Result and inference,
- (2) Action prompt,
- (3) Topic developer and sequencer,
- (4) Connection marker,
- (5) Marker of emergence from incipency,
- (6) Hesitation marker and floor-holding device.

In example (1), the German and the American speaker are talking about tuition fees in the US and how expensive it is for American students to earn a University degree. SB149 shows empathy for the situation that BO03 is describing by saying that he has *such a high weight on your shoulders already* and as a result of that financially difficult situation, she concludes that it is *hard to be academically educated in the US*. The DM *so* in this case is used to **indicate result or inference**.

(1) Result and inference (03SB149BO03; #00:04:30#)

SB149: [...] if you're in debt like a thousands of dollars.
[and],
BO03: [yeah],
SB149: [yo]u have such a high weight on your shoulders already. #00:04:30-3#
.h <D=so> I think it's- it's really hard to be academically,
uhm educated in the US.
BO03: yeah .. ((eh)) it's ... yeah it's very expensive.

In example (2), the DM *so* is used three times in quite a small amount of time. This excerpt is from the beginning of a conversation between an American and a British speaker. The American speaker is trying to open the conversation and get the conversation going by stating *I believe we are supposed to be talking about television shows*. With this utterance, BO15 implicitly wants BI36 to produce the next action, meaning to start talking about TV shows. BI36 however, does not start talking about the topic itself, instead, she agrees that she understood the topic and gives an example what this could mean *what's popular what's not*. With the third *so*, BO15 tries again to open the conversation and she poses a more direct question now to get the conversation about TV shows going *so what's popular in England right now*. BI36 uses *so* (the second *so* in the excerpt) **to mark result**, by saying that from the topic and from her example, she expects an interesting conversation.

(2) Action prompt and topic developer (08BO15BI36; #00:00:03#)

BI36: [[[chuckles]]]
BO15: [u:hm],
... <D=so> I believe we are supposed to be talking about television shows. ((ehh))
BI36: yeah.
what's popular what's not,
<D=so>,
BO15: yes.
BI36: be interesting. [[[chuckles]]]
BO15: [<D=so> what's],
what's uh popular in England right now.

Examples (3.1) and (3.2) both display uses of *so* as **topic developer or sequencer**. In example (3.1.), *so* is used as a topic developer in the opening sequence of the conversation. After both speakers

introduced themselves and asked how they are, SB150 tries to develop a **new topic** by stating: *so, our topic was food, right?*.

(3.1) Topic developer and sequencer – opening (07SB150BI32; #00:00:30#)

SB150: me too thank you ((laughs)),
just had quite a stressful day with university,
but yeah that's fine.
I think you know ((LAUGHS)).
BI32: yeah ..
SB150: .. yeah .. <D=so> our topic was food .. "right"?

In example (3.2.), BI38 and her interlocutor BO18 are talking about different accents and what kind of accents other people like or dislike. They start talking about Northern English and Scottish accents and BI38 states that liking these accents *depends*. She starts a **new topic sequence** telling the story when she went to Huddersfield and people were complaining about the guide's accent. At the end of her story, she comes to the conclusion that understanding and liking accents depends on the people, which she introduces with *so* as a marker of result. She repeats the *I think it depends* hypothesis that she already mentioned right before her story and proves her hypothesis with the argument in her story.

(3.2.) Topic developer and sequencer – storytelling (06BO18BI38; #00:13:30#)

BI38: [...]I think it just depends.
'cause s-is <X=so> many different,
... ways of speaking as well,
<D=so> .. I went to Huddersfield,
uhm a few days ago,
uhm which is just,
kind of northern England-ISH,
u:hm and a lot of people were complaining, #00:13:30-8#
saying I don't know what that man is saying,
and he was,
lecturing,
and telling us about >something<.
uhm but a lot of us didn't understand what he was saying 'cause he had such strong ACcent,
<D=so> I think it really just depen:ds,
°mostly°.

In example (4), *so* is used to **mark connection**. This excerpt is from the closing sequence of the conversation between SB157 and BO11. SB157 wishes BO11 good luck with her studies. BO11 answers

by saying *thanks*. SB157 then also expresses her thanks to BO11 for the talk and *the fun time* they had. BO11 then again also expresses some final wishes to SB157. So, BO11 and SB157 support their respective interlocutor's wishes and thanks by using repetitions. SB157 uses *so* to connect her thoughts across the turns and she connects the wishes and the thank you note in her final closing. All in all, this closing sequence shows very much interconnectivity due to the repetitions.

(4) Marker of connection (06SB157BO11; #00:46:00-5#)

SB157: good luck with your [studies] and everything,

BO11: yeah, {nods}

thanks,

I will .. () for sure {nods},

SB157: <D=so> thanks,

REALLY thanks for this Skype call,

and thanks for the fun time ((chuckles)),

BO11: yeah thank you too,

good luck with everythi:ng, ((chuckles))

In example (5), the participants are talking about their favorite TV series. SB150 says that she is watching *Desperate Housewives* at the moment. She picks up her statement about *Desperate Housewives* again, using *so* as **marker of emergence from incipency**, and refers to the show by explaining why she is watching the show again, i.e. *she is watching it when she is coming home or wants to relax*.

(5) Marker of emergence from incipency (07SB150BI32; #00:40:00#)

BI32: what's your favourite TV show?

SB150: u:hm you know that CHANGES ((laughs)).

BI32: ((laughs)).

SB150: at the moment it's *Desperate Housewives* again.

<D=so> I'm always having one TV show that u:hm,

.. I'm basically WATCHING,

BI32: yeah.

SB150: when I'm coming HO:ME >or when i want to< RELAX,

or just think about something else a:nd, #00:40:00#

BI32: yeah.

SB150: yeah and that's- >at the moment< I'm watching,

Desperate Housewives again ((laughs)).

Example (6.1) and (6.2) display uses of *so* as **floor-holding device and hesitation marker**. This function had been determined in previous research primarily in NNS or ELF contexts (see Buysse 2012, Collet et al. 2021), as example (6.1.) and most of the examples in TaCoCASE underline. However, as example (6.2.) shows, *so* is also used as floor-holding device and hesitation marker by a native speaker (B14), be it in an ELF intercultural setting.

In example (6.1.), the participants are talking about their favorite holidays and SB157 states her favorite holidays are Christmas and Halloween. Halloween is not really a German holiday and SB157 comments that it *is not really that big in Germany unfortunately*. She uses *so* at the end of her utterance to **hold the floor** and in a way also as a **hesitation marker** because she is apparently not sure what other information she wants to add and leaves the utterance open for her interlocutor to jump in.

(6.1.) Hesitation marker and floor-holding device (06SB157BO11; #00:09:30#)

- BO11: ((laughs)) yeah °okay°.
 <so> what's your favorite holiday?
 SB157: m:h o:h I- I think it's a tie,
 between Christmas .. and Halloween probably ((hehe)).
 BO11: a:h very [nice].
 SB157: [u:hm] Christmas for the obvious reasons,
 we already discussed to some extent like the <food>,
 and just the decoration and the good mood every- {moves hands around}
 everyone seems to be in .h,
 but Halloween ...,
 a:h I really love Halloween because I'm really into the spooky .h,
 sort of things and, #00:09:30#
 it's not really that big in Germany unfortunately.
 [<D=so>]-
 BO11: [oh really ah]?
 SB157: not "really" uhm.

(6.2.) Hesitation marker and floor-holding device (05SB124BI14; #00:07:29#)

- BI14: yeah e:r yeah.
 well you're suscept-
 yeah you're susceptible to it,
 when you're in .. in hospital, #00:07:29-9#
 it's like a- it's like a uhm,
 it's a .. yeah it's a virus and it's a RASH.
 SB124: <D=okay>.

BI14: <D=so> .. [it's] ... <D=so> yeah
 SB124: [okay okay yeah .. okay]
 BI14: and u:hm,
 ... <D=so> .. uhm eventually she went to a nursing home.
 .. and she was there for .. four years?
 .. <D=so> .. and it's and it's funny,
 she's from Germany as well= {tilts head}

Example (6.2.) displays various uses of *so* as **hesitation marker** and **floor-holding device**. BI14 is talking about his grandmother who passed away earlier that year and who had to stay at a hospital. He uses *so* in combination with pauses and other hesitation markers such as *uhm* or *yeah*, due to this emotional situation. It is interesting to see that in this example, *so* and *okay* are used in combination, even if they are used by different speakers. This DM combination will now be investigated in greater detail.

DM combination *okay so*

As the results of the collocate analysis (table 32) indicate and as examples (6.2) and (7) demonstrate, we can see that *so* and *okay* are frequently used in **combination**.

(7) DM combination *okay so* (06SB157BO11; #00:24:00#)

SB157: can you still sing the whole song? [...]
 BO11: u:h .. I- I'll you know,
 I'll try (an easy) line of it,
 just 'cause it'll probably make you laugh.
 <D=so> anything [could actually ((laughs))].
 SB157: [o:h <D=okay> ((hehe))].
 BO11: but <D=okay> <D=so>,
 .. I'm not gonna sing I'm just gonna say it,
 if that's <X=okay> ((laughing)).
 <D=okay> oh tannenbaum (German ((0.9))) ,
 ... this is really my (intro) being fun,
 I'm sorry (((laughs))).
 SB157: (((laughs)) (cool)).
 BO11: [u:hm] wie toll sind deine blätter (German ((2.5)))?

For the analysis here, frequency analysis followed by a collocate search (5L and 5R) in AntConc was conducted as a starting point. Again, the tagged version of TaCoCASE was used in order to differentiate between DM and non-DM use of *so* and *okay*. Since it was suggested that *okay* and *so* appear across

speaker turns, the word span was set to 5L and 5R, so that the left and right periphery could be investigated. After that, a qualitative analysis was carried out with special regard to distance between the DM (and what happens in between), identifying uses across speaker turns, uses of separate elements or a combination and these factors were allocated to NS and NNS use.

The frequency of the discourse markers *so* and *okay* in combination will be compared in the different corpora, using TaCoCASE and ViMELF.

TABLE 34: FREQUENCY OF DM COMBINATION *OKAY SO* IN TACOCASE AND VIMELF

	TaCoCASE	ViMELF
raw count	109	190
rel freq (per 1,000 words)	0.77	1.23
same speakers	39%	52%
different speakers	61%	48%
okay so	73%	75%
so okay	27%	25%

As the results in table 34 show, the combination of *so* and *okay* as DM is more frequently used in ViMELF (N=190; 1.23 per 1,000 words) than in TaCoCASE (N= 109; 0.77 per 1,000 words). When examining concordances and the examples more qualitatively, it can be stated that the DM combination is also used across speaker turns, for example when one speaker finished his or her turn with *okay* and the interlocutor started a new turn with *so* (see example 6.2. above). In ViMELF, the DM combination is used equally by the same and by different speakers, i.e. equal number for same turns and across turns. However, in TaCoCASE, the DM combination is more frequently used by different speakers, i.e. across turns. As the collocation analysis (see above) already suggests, there is a sequential preference of *okay* (..) *so* over *so* (..) *okay*, which is the case in both corpora with even almost the same ratio. Some concordances for the DM combination *okay so* are displayed in table 36.

TABLE 35: CONCORDANCES FOR DM COMBINATION *OKAY SO*

No.	ID	Concordance		
1	06SB157BO11	here without, .. her noticing ((hehe)). BO11: ... °ah yeah°. good °(good	okay)° yeah. ...<so>	do you do anything special for your
2	06SB157BO11	t have a bow at hand ((hehe)). #00:09:00# BO11: ((laughs)) yeah °	okay°. < so>	what's your favorite holiday? SB157: m:h
3	07SB149BI31	I've lived here (a while now) ((hehe)) SB149: hehe	okay so	you were born in Birmingham as well? BI31: °>
4	07SB149BI31	no license yet?... BI31: {shakes her head} _<No.> [((heh))] SB149: [	okay,] so	[good luck for that.] BI31: [(It seems, uhm.)]
5	03SB149BO03	standard] BO03: [oh for]. SB149: [stan- standard] one person. BO03:	okay. so	like here we only make giant pizzas, because, ((
6	06BO18BI38	cause we're a fairly RECENT country. so, BI38: oh_	okay so	recent history meaning, .. is there a specific, .. time period?
7	08BO15BI36	excited for Black Panther. .. °personally°. BO15: yes. BI36: [((hehe))] BO15: [	okay], so	I haven't read that comic, #00:50:29-5# in fact

Even from this small excerpt of concordances (see table 35), we see, that the DM combination is used in various ways and that the function of the combination varies as well. Therefore, we will now have a closer look at the function of the DM combination *okay so*.

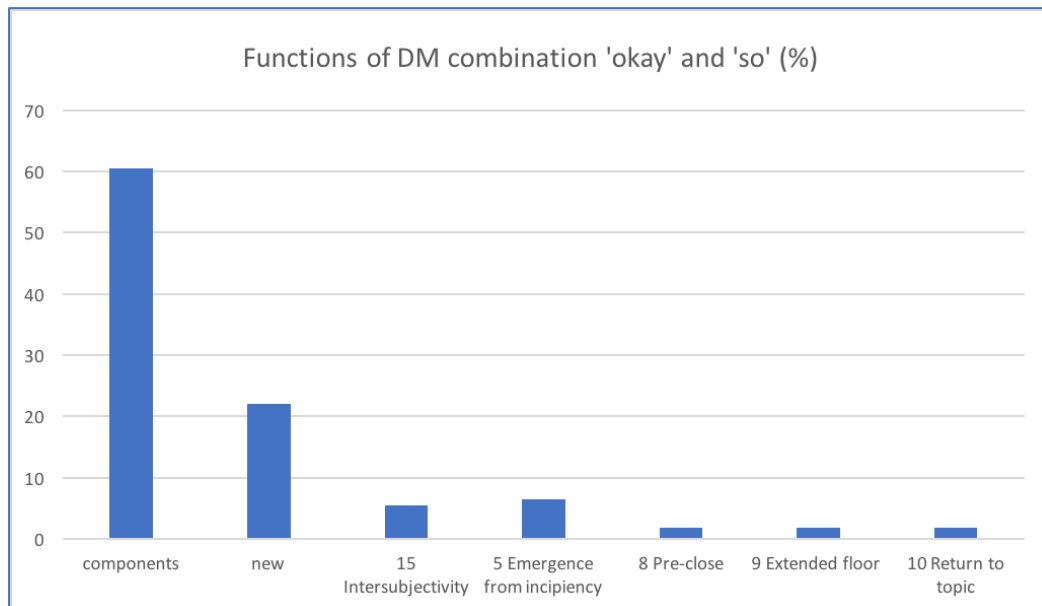
Functions of DM combination *okay so*

As already mentioned, DM frequently co-occur and interact as part of a sequence of two (or more) DM. Typically, this combination of the DM shows differences in functions than the single items (see Koops & Lohmann 2016). In TaCoCASE, the combination of *okay* and *so* appears with an overall frequency of 109. There are several functions that the combination of *so* and *okay* as DM fulfills (see table 36).

TABLE 36: FUNCTIONS OF DM COMBINATION *OKAY SO* IN TACOCASE

combined function of components	N= 66 (61 %)
new function	N= 24 (22%)
one of the functions	N= 19 (17%)

The results in table 36 indicate, that the function of the combination is used in the majority of the cases as a **combined function** of the two components. In 22% of the cases, a **new function** had been determined and in 17% of the cases, the function of the combination results from **one of the functions** of one DM. The distribution for the functions is depicted in greater detail in figure 47.

FIGURE 47: DISTRIBUTION OF FUNCTIONS OF DM COMBINATION *OKAY SO* IN TACOCASE

As far as one of the functions of the DM is concerned, the functions determined were **emergence from incipency** (N=7), which is a function of *so*, as well as **intersubjectivity** (N=6), **pre-close** (N=2), **extended floor** (N=2) and **return to topic** (N=2), which are functions of *okay*.

TABLE 37: FUNCTIONS OF DM COMBINATION *OKAY SO* IN TACOCASE ACROSS AND WITHIN TURNS

	Same speaker (same turn)			Different speakers (across turns)		
Functions	total (same)	<i>okay so</i>	<i>so okay</i>	total (different)	<i>okay so</i>	<i>so okay</i>
Components	14	14	0	52	40	12
New	10	9	1	14	1	13
Intersubjectivity	5	3	2	1	1	0
Em. from incipency	7	7	0	0	0	0
Pre-close	2	2	0	0	0	0
Extended floor	2	1	1	0	0	0
Return to topic	2	2	0	0	0	0

Table 37 displays the individual functions and their respective frequency across and within turns, meaning when used by the same speakers or by different speakers. The DM sequence for the combination is also displayed for the two speaker categories.

What is interesting to see is that the **combined function** from the components is used more frequently by **different speakers across turns** (N=55; 78%) than by the same speaker (N=14; 12%) within a turn. The sequence *okay so* is for both speaker categories pre-dominant (different speakers: 77%, same speakers: 100%).

It can be stated that the overall frequency of the DM combination resulting in **new functions** is slightly higher for different speakers (14; 58%) meaning across turns than for the same speakers (10; 42%) and thus within a turn. What is surprising here is the difference in sequence preference. Over 90% of the new functions were used within a turn (same speakers) in the sequence *okay so*, whereas the preference across turns, when used by different speakers, was *so okay* (different speakers; 13 out of 14 cases; 92%).

Interestingly, the results indicate that if the function of the DM combination results from **one of the functions of one DM**, it appeared mainly within turns, which means that the DM combination was used by the same speaker.

As far as **new functions** of the DM combination *okay so* are concerned, the following potential new functions were identified:

- Abbreviating comprehension,
- Rapport through summary,
- Strengthened intersubjectivity,
- Strengthened topic start.

A distribution of the frequency of these new potential functions is displayed in figure 48.

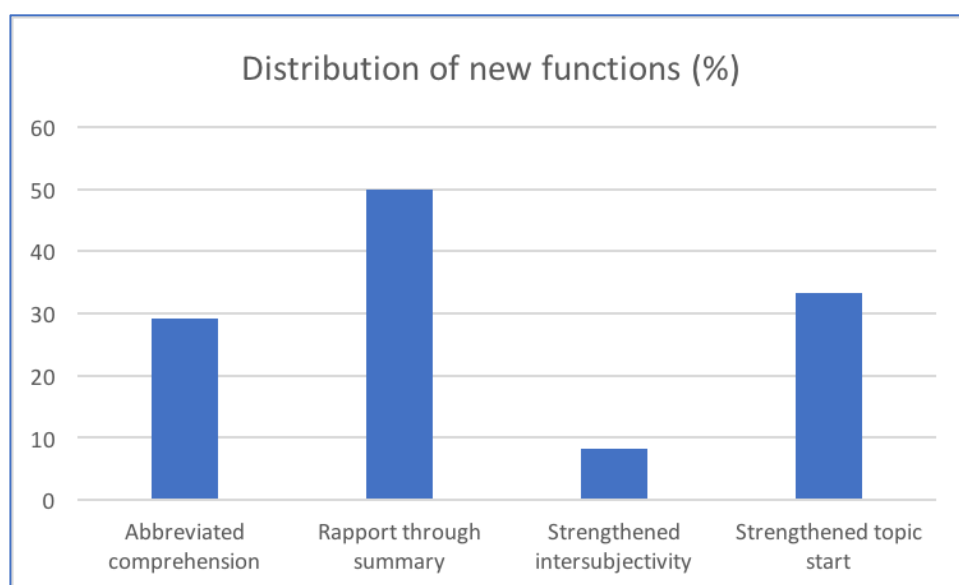


FIGURE 48: DISTRIBUTION OF NEW FUNCTIONS OF DM COMBINATION *OKAY SO* IN TACOCASE

As depicted in figure 48, the new potential function *rapport through summary* is the most frequent among the new functions (N=12; 58%), followed by *strengthened topic start* (N=7; 33%), *abbreviating comprehension* (N=4; 29%) and *strengthened intersubjectivity* (N=1; 8%).

After having looked at new functions of DM in general, I will present some examples from TaCoCASE that illustrate how the speakers used the DM *so* and *okay* with their new functions in the following.

(8) Abbreviating comprehension signal (across speaker turns) (03SB149BO03; #20:00#)

SB149: yeah.

but never mind I can also send you the recording as well .. <D=so>=,

BO03: =<D=okay> .. yeah .. I'll see what I have.

Example (8) is a very common turn final *so*, that is used at the end of an utterance in order to signal that some things are left unsaid. *So* in this example marks connection and *okay* functions as backchannel and together their new function is an **abbreviating comprehension signal**.

(9) Rapport through summary (across speaker turns) (08BO15BI36; #02:30#)

BO15: [yeah].

uh:m,

.. as far as what's popular in America,

... I believe,

.. I actually DON'T have Netflix,

but my aunt does,

<D=so>= ((chuckling))),

BI36: =<D=okay>.

Example (9) is similar to the example before, but in this case it *okay* carries a considerable rapport value. *So* in its individual function would mark result and *okay* is used in its functions of agreement or intersubjectivity. The combination of both contributes to building **rapport** in the sense that the BI36 does not just confirm what BO15 said before, but he actively signals approval.

(10) Strengthened intersubjectivity (same speaker) (01SB151BO05; #10:50#)

BO05: (1.0) uh e:r yeah,

the idea behind,

.. teaching English for me is,

.. to travel ((chuckling)), [...]

.. e:r Asia,

maybe other places eventually.

SB151: .. <D=okay><D=so> that's what you like e:r like best,

In example (10), the DM combination of *so* and *okay* is used by the same speaker. *Okay* is used in its function to create **intersubjectivity** and *so* is used to mark result. SB151 takes over the result from

BO05, but rather than just concluding, it seems that there is a stronger understanding in the sense of establishing intersubjectivity.

(11) Strengthened topic start (same speaker) (06BO18BI38; #04:30#)

- BI38: [>ah<]
 but I don't know what it is though.
 'cause I thought Boxing Day is when you open your presents?
 but I'm- >I don't know<
 BO18: I- ... <D=so>: have you read Harry Potter?
 BI38: yea:h.
 BO18: <D=okay><D=so> I DO remember Harry Potter is-
 somewhere in there Boxing Day is mentioned,
 and looked it up at one point.

In example (11), *okay* in its original function would be a topic starter and *so* fulfills the function of connection marker. Their combination results in a new function that can be described as **strengthened topic start**. It seems that BO18 is indicating that he is about to ask or reveal something that he kind of already announced with his first question and is now trying to get to the point – as if he would say *now let's get real* or *let me get to the good part of my story*.

Differences in use of DM combination among NS and NNS

When the DM combination is used by different speakers, i.e. across turns, there are no differences between either frequency or sequence of use. However, when used by the **same speaker**, there are differences in terms of **frequency and sequence** of the DM combination (see table 38).

TABLE 38: USE OF DM COMBINATION *OKAY SO* ACROSS NS AND NNS (SAME SPEAKER)

	total	okay so	so okay
BI	4	3	1
BO	15	12	3
SB	23	23	0

As the results in table 38 suggest, when used by the same speaker, the German speakers (NNS) use more combinations than the native speakers, i.e. than the American speakers (BO; N=15) and the British speakers (BI; N=4). Interestingly, the **sequence preference is stable across NS and NNS**, which means that all speaker groups preferred the sequence *okay so* over *so okay*. There seems to be a difference in the frequency of use of the DM combination, as NNS seem to use the combination more frequently than NS. No difference among NS and NNS could be identified in terms of DM combination sequence preference.

All in all, the **linguistic analysis** investigated the language use in TaCoCASE in order to provide examples and evidence for the application in the EFL classroom. In using a mixed-methods approach, I combined quantitative and qualitative methods, including corpus linguistics methods such as frequency, collocation, n-gram and concordance line analysis as well as discourse and conversation analysis. The principle that I used is very common in corpus linguistics, which means that first, quantitative methods and in particular frequency is used as a measure to detect language features that need to be investigated in greater detail (see Biber et al. 2004). Additionally, the analysis followed the principle to look at individual words first, then at combination of two words and then at multi-word sequences. In my linguistic analysis, I focused on lexical aspects – words and phrases – as well as pragmatic aspects – discourse markers – as these aspects will be of interest when it comes to teaching linguistic and pragmatic competence and thus communicative competence in the EFL classroom.

The frequency analysis provided the 30 most frequent words in TaCoCASE, which consists mainly of markers of interactivity, such as *I, you, yeah* (as a response token) as well as interactive pronouns and discourse markers (*so, okay, yeah*). These words can as well be found in ViMELF and in other spoken corpora (see O’Keeffe et al. 2007: 12), which suggests that TaCoCASE (and ViMELF) are very representative of spoken corpora. The top 30 word list of TaCoCASE should in general be taught at

school as these words are very frequent and very characteristic of spoken language (O’Keeffe et al. 2007: 12).

The **lexical analysis** was mainly conducted in order to find differences in language use with regard to words (lexis) between native and non-native speakers. In the lexical analysis, I looked at several well-established factors, including vocab profiles (K1, K2, AWL words and off list words) (see *VocabProfiler* by Cobb 2002), lexical density and TTR. The key findings were that there was no big difference as far as the vocab profiles for the speaker groups in TaCoCASE were concerned, which means that the language of NS and NNS differs not very much in TaCoCASE. This is mainly due to the intercultural ELF setting that TaCoCASE is set in (see section 3.1.2.2.4.) and provides evidence that the intercultural speakers seems to be a more suitable model in EFL teaching than native speakers (see discussion in section 2.1.5.2.7.). Small differences were found in terms of lexical density and TTR. The findings indicate that NS used more content words and thus more variation in their utterances compared to the non-native speakers, which is rather due to personal preference than differences in proficiency.

As far as combination of words, in this case **collocates**, are concerned, I looked at *really*, which was found to be a very frequent adjective in TaCoCASE. Collocates of *really* that were found are other adjectives such as *nice*, *big*, *cool* and *interesting* as well as verbs like *enjoy* or *know*. *Really* is used in TaCoCASE in combination with adjectives and verbs, usually as modifier, intensifier or booster in order to express emphasis or truthfulness (see Biber et al. 2002; Pérez-Paredes & Bueno-Alastuey 2019).

As far as n-grams are concerned, the most frequently used n-grams in TaCoCASE were summarized in tables 26 and 27 (3-grams and 4-grams) and the most frequent 3-gram was *I don’t know*. A common frequent pattern that had been identified, which is also very common in previous research (Biber et al. 2002: 446) can be described as follows: first person singular pronoun *I* + negative auxiliary verb (*don’t* as contraction of *do not*) + main verb (*know*) + complement. The complement was either introduced by interrogative pronouns such as *what* and *how* like *I don’t know what it is* or *I don’t know what was wrong with it*, or by the conjunction *if*, followed by an if-clause, like for example *I don’t know*

if I actually would want to. Most of the lexical bundles found in TaCoCASE are stance bundles (Biber 2004, Simpson-Vlach & Ellis 2010) and referential expressions (see Biber et al. 2004; Simpson-Vlach & Ellis 2010). Among the stance bundles, I identified personal epistemic bundles, such as *I don't know if/what/how*, which was the most frequent n-gram in TaCoCASE as well as the most frequent negative collocation in COCA and the BNC (Baumgarten & House 2010: 1186), and *I (don't) think* (see Biber et al. (2004). *I don't know* is a personal epistemic stance bundle (Biber et al. 2004: 389) used to express uncertainty and insufficient knowledge (see Tsui 1991; Baumgarten and House 2010). *I (don't) think* is another frequently used stance bundle, which was used to express “possibility, but lack of certainty” (Biber et al. 2004: 389). Another stance bundle is *it's kind of*, which is used as a hedge to express ambiguity and probability (see also AFL, Simpson-Vlach & Ellis 2010). Only one referential expression had been identified in TaCoCASE, *a lot of*, which was used to specifying quantity (see Biber et al. 2004: 395, see also Simpson-Vlach & Ellis 2010).

As far as the **pragmatic analysis** is concerned, I looked at the most frequent **discourse marker so** and analyzed it with regard to frequency, collocates and its functions. Frequent collocates of *so* that could be found in TaCoCASE are *many*, *much* and *okay*. Functions of *so* that had been established in previous research can as well be found in TaCoCASE, such as result and inference, action prompt, topic developer and sequencer, connection marker, marker of emergence from incipency, as well as hesitation marker and floor-holding device (Schiffrin 1997; Buysse 2012; House 2013; Redeker 2006; Fischer 2006; Bolden 2006; Collet et al. 2021). *So* co-occurred with other discourse markers, such as *okay*, which made me analyze this combination further. The results suggest that the combination *okay so* is preferred over *so okay* in TaCoCASE, even across speaker turns. The functions of the DM combination is in most of the cases a combined function (61%) and the combination is only to a smaller extent with new functions or one of the functions of the components. The new functions of the combination *okay so* that were found are: Abbreviating comprehension, rapport through summary, strengthened intersubjectivity and strengthened topic start. In terms of differences across native and non-native speakers, the results indicated that there was a difference in DM combination use in the

sense that NNS seem to use the combination more frequently than NS. However, there was no difference found among NS and NNS in terms of sequence preference.

The results from the analysis provide relevant findings and information for application in EFL teaching. Not only the findings, but also the examples that were discussed, could be used in EFL teaching. In terms of frequency and lexis, the analysis provided frequent words, collocations and lexical bundles, which means words and phrases that are frequently used in the corpus and which are authentic and relevant. These words and phrases as well as the examples discussed should be used in the EFL classroom. Additionally, most of the analysis were conducted with corpus tools, which could also be used in a DDL activity and thus the learners could conduct similar analysis themselves and come up with similar findings. This frequency-based approach provides also proof and guidance in the selection of vocabulary and it provides answers to the question what words and phrases should be taught in the EFL classroom. As far as pragmatic elements and in particular discourse markers are concerned, the results provide implications on how discourse markers are used with regard to frequency, form and functions, as well as combinations of discourse markers. Given the fact that discourse markers usually provide difficulties for learners, the findings and the examples that were discussed here, could be very helpful for learners.

4.2. (INTER)CULTURAL ANALYSIS

Talking about **culture and cultural aspects** is the main characteristic of intercultural communication. Mediation as well as exchanging (new) cultural concepts is one key element of intercultural competence. As TaCoCASE is set in an ELF intercultural environment (see section 3.1.2.2.4.), it allows research not only from a linguistic perspective, but also in the field of intercultural communication. The benefits of corpora for language teaching and the investigation of corpora with regard to linguistic aspects has been researched and acknowledged widely (see for example O’Keeffe et al. 2007, McEnery & Hardie 2012, Hunston 2002, Römer 2011, Mukherjee 2002; see also section 2.2.2.4.). Using corpora for **investigating intercultural communication** and **teaching intercultural competence**, however, is not very common and a “novel approach” (Fahey Palma 2022: 116). According to Fahey Palma (2002), it can be “transformational to students’ intercultural development” (Fahey Palma 2022: 116). Using a corpus for analysis of intercultural communication is one key benefit of TaCoCASE and one key contribution of this dissertation. In this chapter, I will investigate what the participants talk about in their intercultural encounters with special regard to **cultural aspects**. In a second step, I will analyze how the participants talk about cultural aspects in terms of **discourse structure** and **strategies to explain (new) cultural concepts**. This intercultural analysis is designed to set the basis and provide findings for the application of the corpus in the EFL classroom when it comes to establishing intercultural (communicative) competence.

As far as the **methodology** is concerned, there are four main well-established approaches for researching intercultural communication in education (see Godley 2012: 461): Ethnography (see for example Hymes 1972, Pike 1967), (critical) discourse analysis (see for example Fairclough 1995, Gee 2007), conversation analysis (Gumperz 1982; Sacks et al. 1974, Kasper 2006, Mori 2002) and systemic functional analysis (Halliday 1973). While ethnographers investigate intercultural communication from the speakers’ perspective and their view of cultural practices and norms, systemic functional linguists look at the combination of function and meaning of words that are used, which means subject to analysis is how grammar and lexis influence meaning. These two approaches are either very subjective

or very systematic, which is not very useful for the purpose of this study and the application in a teaching context. (Critical) discourse analysis and conversation analysis focus on the conversation and discourse itself and how people create meaning and organize discourse. Since the focus in this dissertation is on the combination of communicative and intercultural competence and thus the combination of linguistic, pragmatic and discourse as well as intercultural elements, **discourse analysis** and **conversation analysis** are more suitable with regard to the aim of this study.

I am using a **mixed methods approach** in combining corpus linguistics with discourse analysis and conversation analysis in order to investigate cultural issues and practices (see Baker 2006, McEnery & Hardie 2012; see also section 3.2.). In a combined discourse and conversation analytical approach, I will investigate how the speakers interact, how they organize discourse in order to successfully communicate, what discourse strategies they use to explain (new) cultural concepts and how they create and negotiate meaning in interaction. Not only the discourse structure and elements will be of interest, but also what the people said, when and why (Seedhouse 2005) as well as multimodal features will be taken into account when analyzing the conversations. This mixed methods and multimodal approach, including quantitative and qualitative elements, promises to reveal more fruitful results than one single approach (Dörnyei 2007).

The **negotiation of meaning and cultural concepts** plays an important role in intercultural communication. While research and teaching intercultural communication has for a long time focused on misunderstandings and failure of communication due to cultural differences (see Lüsebrink 2016), I take a different approach that is usually carried out in ELF research. I look at **communicative success** (Wagner & Firth 1997) and on the co-construction and negotiation of meaning as a collaborative effort, which contributes to mutual understanding (Baker 2015: 34; see also Hua 2015: 85). ELF is “intercultural in nature [...] in that it concerns communication that takes place among speakers from various linguacultural backgrounds” (Cogo & Dewey 2012: 23). I agree with Hua (2016: 8) that culture can only be interpreted as “the result of observed and contextualized practices emerging in communication”, focusing on the discursively constructed way of negotiating meaning. In terms of

intercultural communication, I follow Baker (2015) that cultural differences and similarities, as discussed in “*intercultural communication*” can be “changeable and negotiable at the centre of intercultural communication, hence seeing cultures as heterogeneous and dynamic entities that emerge from communication” (Baker 2015: 21). In ELF intercultural settings like in TaCoCASE (see section 3.1.2.2.4.), negotiation and understanding is key as “shared knowledge and common ground cannot be assumed” (Cogo & Dewey 2012: 44). Hua defines negotiation of meaning – which she differentiates from negotiation as in business contexts – as a “mechanism that enables participants in intercultural and lingua franca communication to employ, mobilise, or manipulate diverse resources to achieve their goals of interaction” (Hua 2015: 64). In a similar manner, understanding is a process of “building common ground or joint knowledge, rather than taking these for granted (Dewey 2012: 45; see also Pitzl 2010). Negotiation of meaning and understanding thus can only be achieved through discursive interaction, which then contributes to creating common ground between the speakers and leads to successful communication.

I combine the **ELF and intercultural approach** as both “explor[e] how participants with different linguistic-cultural backgrounds negotiate meaning in intercultural encounters” (Hua 2016: 173). TaCoCASE is very well suited for this approach as it is a corpus that is set in an ELF intercultural environment (see section 3.1.2.2.4.). Thus, the aim of this intercultural analysis is to find out how successful communication is achieved when meaning and cultural practices are negotiated. I look at ELF intercultural communication focusing on successful communication – not on misunderstandings – and not only at traditional *intercultural* aspects such as similarities and differences between cultures, but also at transcultural and hybrid aspects of culture (see section 2.1.5.2.).

As far as **intercultural communication** is concerned, there are several factors that can be drawn upon in the analysis (Müller-Jacquier 1999, cited in Lüsebrink 2016: 52) and topics as well as discourse organization are two of them. I follow this analysis scheme and established the research questions based on topics and discourse organization in intercultural communication.

The following **research questions** guide through the analysis in this chapter.

Research Questions:

- RQ5: What aspects of culture can be found in TaCoCASE?
- RQ6: How do the participants talk about cultural aspects?

The research questions are based on methodological considerations as well as requirements for the application of the corpus in an EFL teaching context. Methodological considerations include the combination of corpus linguistics with discourse analysis and conversation analysis and with it the investigation of how people interact – from a discourse point of view – and how they negotiate meaning – from a cultural point of view – based on conversations taken from a corpus of ELF intercultural encounters (TaCoCASE). Requirements and application in an EFL teaching context is dependent on intercultural competence and what aspects of intercultural (communicative) competence the learners should acquire.

I follow Michael Byram's **ICC model** (see Byram 1997, 2021), which is well-established in EFL teaching and the intercultural learning and intercultural sensitivity model (Bennett 1993; see also Barmeyer 2000) (for a detailed explanation see section 2.1.5.2.). In Byram's ICC model, the **five *savoirs*** are key to establishing intercultural (communicative) competence: (1) knowledge, (2) attitudes, (3) skills of interpreting and relating, (4) skills of discovery and interaction and (5) critical cultural awareness. The research questions focus on **cultural topics** – what the participants talk about – and how the participants talk about cultural aspects. Hence, they draw on three of the four *savoirs*, namely knowledge, skills of interpreting and relating and skills of discovery and interaction. To some extent, they also draw on attitudes and critical cultural awareness, even though these are elements that should be established by the learners on a personal level in intercultural encounters. Knowledge, skills of interpreting and relating as well as skills of discovery and interaction could be established from looking at (authentic) examples – as it is intended here with examples from TaCoCASE. Knowledge is concerned with knowledge of social groups and practices (Byram 1997: 51), skills of interpreting and relating means finding differences and similarities between own and other cultures and skills of

discovery and interaction is concerned with acquiring knowledge of cultural practices and being able to apply knowledge, skills and attitudes in (real-time) interaction (Byram 1997: 52-53). This means acquiring knowledge about cultures and cultural practices, as well as relating and interpreting cultures in a sense of comparing differences and similarities between own and other cultures as well as being able to interact in intercultural communication are aspects of intercultural communicative competence, which are not only suggested by Byram's ICC model, but which are taken as a basis for the analysis and the research questions in this section. As far as intercultural learning is concerned, Bennett suggests that with growing experience of cultural differences and action competence, learners establish intercultural sensitivity (see Bennett 1993: 182) and the ability to interact in intercultural situations. Combining these models means that establishing knowledge and skills as well as intercultural sensitivity leads to intercultural communicative competence.

Furthermore, there is a strong demand in **intercultural learning** in the EFL classroom to shift from "being to doing culture" (see Kramsch & Hua 2016), which calls for a **discursive perspective** on ICC (see Hallet 2008; Altmeyer 2017; König et al. 2022). Using examples from TaCoCASE and looking together with the learners at how the speakers discursively construct meaning and negotiate cultural concepts corresponds to this demand and this new approach in intercultural learning in the EFL classroom. It follows thus a new and modern way of establishing ICC and provides material that could be used in the EFL classroom, which is quite rare as the approach is very new and not many material exists. In a further step, the learners should learn to act interculturally competent in intercultural encounters themselves, but as a starting point and in order to teach mediation strategies and how to explain cultural concepts, this approach is very promising and should be integrated in intercultural learning and EFL teaching.

The research questions were based on these assumptions and on the combination of the models described above in order to establish intercultural communicative competence and intercultural sensitivity. **RQ5** is mainly concerned with **knowledge about social groups, practices and norms**. It will be investigated what cultural aspects the participants talk about in terms of food, customs and

traditions and popular culture and to what extent these aspects provide information about their own culture or about the interlocutor's culture.

RQ6 focuses on **skills of interpreting and relating** as well as **skills of discovery and interaction**, which means that it will be investigated how the participants compare cultural aspects and how they interact with each other to negotiate meaning. This includes looking at **(cultural) differences and similarities**, when relating and interpreting cultures. It will also be analyzed how the participants explained (new) cultural concepts in terms of discourse elements and strategies.

These findings contribute to understanding how the participants talk about cultural aspects, how they interact and how they negotiate meaning of cultural concepts. It provides information about knowledge and skills about the own and other cultures and paves the way towards intercultural competence and intercultural sensitivity. Thus, the (inter)cultural analysis provides results that can be used as a basis for the application of TaCoCASE in the EFL classroom (see chapter 5).

4.2.1. CULTURAL ASPECTS IN TaCoCASE

RQ5: What aspects of culture can be found in TaCoCASE?

The participants were provided with topic prompts for their recordings (see section 3.1.2.2.2.) and some topic prompts target cultural topics. Due to this fact and due to the fact that the recordings represent intercultural encounters, it can be expected that cultural topics play an important role in the data. Almost $\frac{3}{4}$ of the conversations covered cultural topics, which supports this hypothesis. The topics I was especially interested in and which were represented by the topic prompts are: **food, customs and traditions** and **popular culture**.

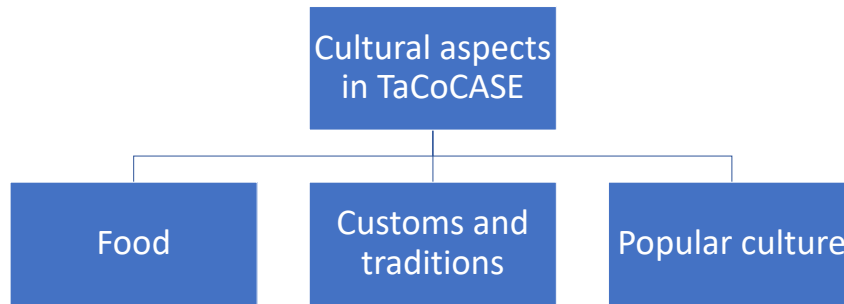


FIGURE 49: CULTURAL ASPECTS IN TACOCASE

The participants were not restricted to the provided topics and they could choose the topics they wanted to cover during the conversations themselves. This means that cultural topics could arise in every conversation. The topic prompts associated with cultural aspects are the following:

6. “How do you celebrate?” – Cultures and traditions
7. “Let’s talk about food” – Eating habits and preferences – what do you like to eat?
8. “What’s on?” Popular Culture – Talk about your favorite TV show

As far as these topics are concerned, they all clearly belong to culture and describe aspects of culture.

Food is usually connected to culture and lifestyle. It depends on different cultures as it can be seen as a “bridge between nature and culture” (Fischler 1988, cited in Germov & Williams 2008: 1). Apart from aspects of food belonging to specific cultures, food is usually also related to rituals (see Hofstede 1993), as some rituals include meals and it becomes a tradition to serve certain dishes at certain festivities or rituals (see Rüdiger & Mühleisen 2020: 2). Food is also used to construct identities in a sense that “[e]very coherent social group has its own unique foodways” (Counihan 1999: 6), which makes it a very culture-specific phenomenon. Rüdiger and Mühleisen (2020), who looked at the linguistics of food and published the book *Talking about food – the social and the global in eating communities*, state that “[t]alking about food is an everyday activity” (Rüdiger & Mühleisen 2020: 1). In their edited volume, they and many contributing authors discuss the relation of food and tradition, health aspects, morality (with regard to self-imposed dietary restrictions such as veganism) (see Dreschler 2020) as

well as food and religion (Rüdiger & Mühleisen 2020: 2). It should be mentioned here that food can be very identity-forming and culture-specific, even though with globalization and the rise of social media, food trends become global phenomena and can also be transcultural phenomena.

Customs and traditions also clearly belong to culture and are defining aspects of a culture (see Hofstede's cultural dimensions 1993: 22; see also section 2.1.5.2.). As far as **popular culture** is concerned, there is not one clear cut definition what popular culture is. Williams (1993) described popular culture as "Culture actually made by the people for themselves" and it is "well liked by many people" (Williams 1993: 273; cited in Storey 2006:5). Storey (2006) came up with **six aspects** to what the concept of popular culture is... (Storey 2006: 5ff):

1. ... "simply a culture that is widely favored or well-liked by many people" (Storey 2006: 5)
2. ... "what is left over after we have decided what is high culture", i.e. "texts and practices that fail to meet the required standards to qualify as high culture". (Storey 2006: 6)
3. ... a "mass culture", i.e. it is "mass-produced for mass consumption" (Storey 2006: 8).
4. ... "the culture that originates from 'the people'" in a sense that it is a "folk culture" (Storey 2006: 9)
5. ... as a "terrain of exchange and negotiation" between "the imposed culture of the mass culture theorists" and the "culture of 'the people'" (Storey 2006: 10)
6. ... should be seen in light of "recent thinking around the debate on postmodernism", where the distinction between high and popular culture is no longer recognized and elitism is put to an end. (Storey 2006: 12)

It becomes clear that "different theoretical perspectives have tended to focus on particular areas of the popular cultural landscape" (Storey 2006: 14). On the one hand, they focused on "the study of texts", including for example television and pop music and others on "lived cultures or practices", such as for example youth subcultures (Storey 2006: 14). Discussions about postmodern or popular culture include both, "cultural texts and practices, for example television, music video, advertising, film, pop music, fashion" (Storey 2006: 197). Food and customs and traditions thus contribute to popular culture as well (Rüdiger & Mühleisen 2020; Santino 1996; Rutagand 2024) . Popular culture has a global nature and the boundaries are fluid, which means that it could be regarded as transnational phenomenon. Astrid Fellner et al. (2018) state that (pop) cultures are "on the move" and that cultural mobility drives cultural exchange, which means that "cultural exchange permits concepts to be exchanged and shared

by different societies” (Fellner et al. 2018: 7). Pop Cultures move from West to East and back, which ends up in the formation of a “densely connected global network” (Fellner et al. 2018: 7). In my data, for example, the participants usually talk about typical food from a certain country or a typical tradition or festivity that is celebrated in a certain country. However, with popular culture, countries and national borders are not in the center of the discussion anymore and TV series for example are usually not put into categories from which country they stem from. The reason for this is probably that American pop culture has influenced European countries and that most of the TV series and shows people watch are provided on global (American) streaming platforms.

When it comes to **language teaching**, pop culture and its artifacts should be integrated in language teaching and learning as related topics can be found in language curricula. Popular culture plays a “highly relevant role in the lives of many learners” (Werner 2022: 161). Learners watch TV series or listen to music that is usually in English and thus, these artifacts should be used in the classroom as well. Werner and Tegge (2020) argue that pop culture and its artifacts are a good starting point for developing language-related competences in the EFL classroom and provide interesting insights on how pop culture can be used in language teaching and learning in their book *Pop Culture in Language Education – Theory, Research, Practice*. I agree with Werner and Tegge and I would suggest to **integrate aspects of (popular) culture in the EFL classroom** for two reasons: (1) It is close to the learners’ life and world experience, which makes it very motivating for the learners to talk about these aspects (see Deci & Ryan 1993 on motivation). If learners watch movies or TV series in their free time and discuss these in the EFL classroom in a foreign language, this contributes to communicative competence. It combines the world inside the classroom with the world outside the classroom and thus makes EFL teaching more authentic and closer to the real world. (2) In discussing TV series or aspects of popular culture, it raises the learners’ awareness of transcultural and hybrid aspects of culture and it contributes to intercultural competence when they discuss TV series in a (critical) cultural sense. Hence, integrating elements of popular culture into the EFL classroom contributes to all five *savoirs* in

Byram's model of intercultural communicative competence (see Byram 1997, 2021) and to intercultural sensitivity (see Barmeyer 2000; Bennett 1993).

Thus, in TaCoCASE, when looking at food and culture and traditions, it seems that the participants talk mainly about culture-specific aspects, which are closely related to national cultures. When talking about TV series or popular culture, it becomes clear that the participants see these aspects and topics as transcultural or hybrid phenomena. As far as teaching and establishing intercultural competence is concerned, it is important to raise awareness for cultural aspects in both senses – from a **national perspective and a transcultural perspective**. The selection of topics – food, customs and traditions and popular culture – depicts these perspectives and the topics as well as how the participants talked about these topics will be analyzed in the following.

Analysis

Eight out of the 15 conversations had topic prompts with cultural aspects: Two conversations were allocated topic 6 (culture and traditions), three conversations were allocated topic 7 (food) and three conversations were allocated topic 8 (popular culture). The following table depicts the distribution of cultural topics – food, culture and traditions and popular culture – across the conversations.

TABLE 39: CULTURAL ASPECTS IN THE CONVERSATIONS

Topic	Conversation ID	Food	Culture & Traditions	Popular Culture
1	01SB151BO05			
2	02SB152BO06			
2	02BI33BO10			
3	03SB156BI24		x	x
3	03SB149BO03	x		x
4	04BO16BI40			
5	05SB124BI14		x	x
6	06SB157BO11	x	x	x
6	06BO18BI38		x	
7	07SB149BI31	x		x
7	07SB150BI32		x	x
7	07SB156BO10	x	x	
8	08SB147BI29		x	
8	08BO15BI36		x	x
8	08BO18BI41			x

It is interesting to see that cultural aspects such as food, customs and traditions and popular culture, were discussed in more conversations than could be expected from the topic prompts – in fact 11 conversations instead of eight. Only four conversations out of 15 do not cover cultural topics such as food, customs and traditions and TV shows. This means that **almost 3/4 of the conversations in TaCoCASE are concerned with cultural aspects**. All conversations which should cover cultural topics did cover (at least) one cultural topic, usually even more than one. As the participants were free to choose the topic they would like to talk about – they had topic prompts but the topic prompts were not mandatory – the participants chose cultural topics voluntarily and consciously. Cultural topics seem to be good conversation topics for first encounters as the conversations were very successful and the participants typically covered more than one cultural topic. In talking about these cultural aspects, the participants showed cultural openness, curiosity and awareness towards their interlocutor and their interlocutor's culture. The cultural topics help creating rapport (see Brunner 2021; Diemer & Brunner 2020) and building common ground between the participants and this is usually accompanied by laughter (see Schmidt 2020) in TaCoCASE. Talking about cultural topics also allows the participants to position themselves, with regard to their own culture and with regard to a different culture, which

helps the speakers to create identity. This shows very well that TaCoCASE is set in an ELF intercultural environment and that cultural topics play an important role in intercultural communication and in establishing common ground and rapport.

In order to find words that are related to food in TaCoCASE, I carried out a **semantic analysis**, looking at words and their meaning. I followed a corpus-based approach to semantic analysis and used the UCREL Semantic Analysis System (USAS)²¹, a semantic tagger that finds words related to food and adds tags to these words. The tagset allows searches for the **tag_F1**, which is characterized in the tagset as **food**. Unfortunately, UCREL does not provide tags for the two other cultural topics such as customs and traditions or aspects of popular culture, which is why we can only look at food in this semantic analysis.

The results show **540 words** that are **semantically related to food**, which means that words related to food appear roughly 4 times per thousand words in TaCoCASE. Some examples that I found are: food, eating, cookies, turkey, meat, candy, picnic, mashed potatoes, starter, rice, pizza, beef, grocery store, recipe. These words can be put in different categories:

- Words that describe **food items**: cookies, turkey, meat, mashed potatoes, rice, pizza
- **Verbs** that are used in context with food: eat, bake
- Words that describe **locations** in context with food: grocery store, restaurant
- **Others**: picnic, starter, meal, recipe

This list gives us already an idea what the participants talk about and provides a good semantic picture of the words that are related with food. Rüdiger and Mühleisen, who investigated the linguistics of food, found that food discourse is usually concerned with “what, where, and how to eat” (Rüdiger & Mühleisen 2020: 2). The categories that I established are very similar to what Rüdiger and Mühleisen describe as the food items represent the “what”, locations of food represent the “where” and verbs

²¹ <https://ucrel.lancs.ac.uk/usas/>

and other categories belong to the “how” to eat. It can be stated that the conversations about food in TaCoCASE follow established topics and elements of food discourse (see Rüdiger & Mühleisen 2020).

I conducted a more **qualitative analysis** in order to specify the topics the participants talked about. The conversations were analyzed with regard to the cultural topics – food, customs and traditions and popular culture – and the topics the participants covered was extracted from listening to the conversations and noting down the different topics. In a next step, I clustered the topics and categorized them. The following figure provides a more detailed overview of the topics that were found in the conversations and that the participants talked about.

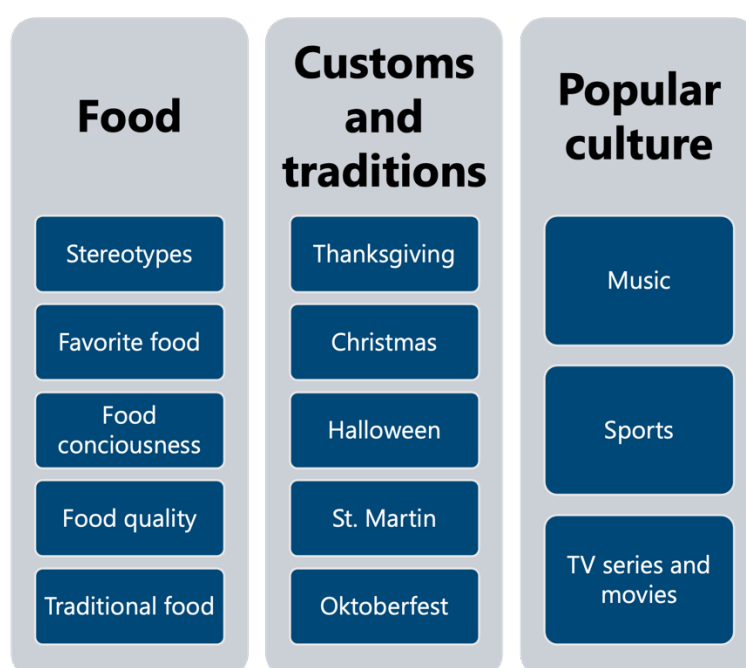


FIGURE 50: EXAMPLE TOPICS FROM CONVERSATIONS ABOUT CULTURAL ASPECTS

These topics and some examples will be explained and analyzed in more detail in the following.

4.2.1.1. FOOD

In terms of **food**, participants in TaCoCASE talked about **stereotypical food**, for example *fish and chips* as a stereotypical British food (see example 12), *burgers* as a very stereotypical American fast food (07SB156BO10; #15:00#) and *Weisswurst* or *Brezel* which is very common in Germany (07SB156BO10; #27:25#).

(12) Stereotypes – fish and chips (07BI31SB149; #02:00#)

- SB149: ((snuffles))..u:hm..
but uh also in terms of food u:hm {lifts head slightly & raises eyebrows}
it's where I lea-I learnt to love **fish and chips**. (1.2)
- BI31: **FISH AND CHIPS** ((laughing)). {leans forward}
.h not the **stereotype**, isn't it?
u[:hm, that..it's-]
- SB149: [yeah it is (.)] {moves hand to chin}
- BI31: English people with their fish and chips. ((chuckles))
- SB149: yeah but it's.. it's amazing.
I can - I can- {moves hand in circle}
if it's not uh.. if it wouldn't be a stereotype,
I could fully understand why uh [people] like it.
- BI31: [mhm.] {smiles}
.h °Yeah.° {looks up}
there is literally u:hm
"outside my house is a fish and chips (shop)" so. #00:02:33-0#

In this example, the German speaker SB149 states that during his stay abroad in England, he *learnt to love fish and chips*. The expression *learnt to love* implicitly reveals that it was a process for him: he presumably did not like it at the beginning, but liked it very much (*loved it*) in the end. BI31 makes fun of him by repeating the words *fish and chips* very loudly while laughing at the same time- This reaction signals creating rapport while at the same time signaling understanding and approval to elaborate on the stereotype. The British speaker also uses irony in her rhetorical question *not the stereotype, isn't it*. It is very interesting that BI31 uses the word *their* (*fish and chips*), positioning herself outside the British culture and referring to what other people say about the British people by mocking it, which is indicated by laughter. SB149 defends himself and the British people by claiming that he really liked it and that it is really tasty (*amazing*), probably to prevent a possible face-loss and to relativize the

stereotype in order to create common ground. BI31 affirms the stereotype by saying that *fish and chips* is eaten everywhere in England and that a fish and chip shop is right outside her house. In this example, it becomes very clear that the participants are creating rapport by laughing while talking about stereotypes. Stereotypes are used in a positive way and help to create common ground as both participants can relate to it. This is very unusual as stereotypes are usually used in a negative way and they are also usually depicted in textbooks with a negative connotation. This example helps create a new, positive image of how stereotypes help to create common ground and can actually be used to establish intercultural competence.

In 03SB149BO03, 07SB149BI31 and 07SB156BO10, the participants talk about their **favorite food** and what they like or dislike. In 07SB149BI31, it is used as a starting point before diving into a deeper conversation about food. Asking people what their favorite food is or what they like to eat is a good “small talk” topic as everyone can say something about it and is very well suited as a topic starter for a food conversation. It has an evaluative and affective notion as people usually express like or dislike and the interlocutor can also relate to the topic and usually also to the food or the discussed item. The affective notion is also something which is important in terms of intercultural learning, as this notion is usually disregarded in EFL teaching (see König et al. 2022: 8).

Food consciousness and **food quality** are big issues in conversation 07SB149BI31 and 07SB156BO10. In 07SB149BI31, the participants talk about healthy food and end up in a discussion about organic food and how the participants prepare their own food (#9:18-10:55#). They also talk about food consciousness and about Halal food, as BI31 is Muslim and only eats Halal food. Lifestyle as well as religious aspects of food seem to be topics to talk about in conversations about food (see also Rüdiger & Mühleisen 2020; see also Drescher 2020).

In 07SB156BO10, SB156 explained that he turned into a vegan and how he came to that decision. Milk alternatives (#9:18#), animal suffering (#12:25#), the conditions under which meat is produced or living independently and growing food in an own garden (#20:40#) are topics the participants covered in

their conversation. Talking about food quality, the participants put food quality in relation to money and how food becomes more expensive the better the quality is (07SB156BO10 #17:35#; 07SB149BI31; #12:00#).

In 07SB156BO10, the participants talk about **traditional food** at the *Oktoberfest*, which is a German festivity that takes place in Munich every September and is considered to be very traditional and typical for the Bavarian region. BO10 knows *Oktoberfest* and explains that there are festivities that draw on *Oktoberfest* in America as well. At *Oktoberfest*, traditional (Southern) German food is served such as “Brezel”, “Weisswurst” and beer. An excerpt of the conversation is depicted in example 13.

(13) Traditional food – Brezel, Weisswurst (07SB156BO10; #33:00#)

SB156: (2.0) and uhm ... you serve it with uhm .. sweet honey? {tilts head}
BO10: mhm. {nods}
SB156: **white sausage.**
BO10: OOH.
SB156: =[you know] Germans call it **Weisswurst ((German (0.3)))**.
BO10: mhm.
SB156: Weisswurst ((German (0.3)))
BO10: OH >yeah – yeah – yeah< okay.
[it's a the same–]=
SB156: =[what's the word in English]
BO10: I think it's the same word in English,
just butchered,
we say **bratwurst.**
SB156: .. OH okay. {nods}
BO10: yeah. #00:33:30#
SB156: and it's – it's uhm .. it's served,
in Bavaria and especially during uh the Oktoberfest ((German (0.4))).
BO10: mhm okay.
SB156: **it's typically Bavarian.**

The participants were talking about *Brezel*, a type of baked pastry, shaped in a knot that is made from lye dough and usually has salt on it. It is also known in the U.S, even though in the U.S., they are usually soft, not made with lye dough and do not have salt. The participants have an argument about the correct spelling and SB156 demonstrates his expert status by explaining that it is usually eaten with sweet honey (note: he probably means sweet mustard). He continues to explain that *Weisswurst* is usually eaten together with a *Brezel* and BO10 confirms that she knows what *Weisswurst* is and she

relates to the American version, called *bratwurst* in America. BO10 signals familiarity with the concept of *Brezel* and *Weisswurst* and by using the word *bratwurst*, she is demonstrating her co-expert status as far as the American equivalent of the German concept is concerned. SB156 then concludes that the items mentioned before are usually eaten at *Oktoberfest* and that it is a typical Bavarian or (Southern) German tradition, drawing on the local and regional cuisine in Southern Germany and Bavaria in particular. Again, he expresses his expert knowledge and relates the food item to a traditional festivity. As this example 13 demonstrated, food and customs and tradition are usually intertwined and food plays an important role when it comes to customs or traditions (Rüdiger & Mühleisen 2020: 1; Rutagand 2024; Santino 1996).

The topics that were discussed in terms of food can also be found in the research on food discourse (see Rüdiger & Mühleisen 2020). Brunner and Diemer (2020), for example, found participants talking about stereotypes in their data, when investigating expert identity in ViMELF (Diemer & Brunner 2020). They also found “evaluative statements about food” (Diemer & Brunner 2020: 177), which is very similar to what I found in TaCoCASE when the participants talked about their favorite food. Preference of certain food is usually used to create individual and group identity (Rüdiger & Mühleisen 2020: 2). Food consciousness and food quality are topics that came up in a conversation about veganism as a lifestyle in TaCoCASE and can thus be regarded as lifestyle issues and morality aspects (Rüdiger & Mühleisen 2020, Dreschler 2020). Dreschler (2020), for example, looked at veganism and vegetarianism and moral aspects of these self-imposed diets. Conversations about traditional food in TaCoCASE are usually related to local cuisine or regional cuisine, which are also an important aspect when talking about food (see Rüdiger & Mühleisen 2020; Rutagand 2024; Santino 1996). It can be stated that the topics that were discussed in terms of food in TaCoCASE, such as stereotypes, favorite food, food quality and food consciousness as well as traditional food, can also be found in established research (see for example Rüdiger & Mühleisen 2020, Diemer & Brunner 2020, Dreschler 2020).

4.2.1.2. CUSTOMS AND TRADITIONS

As far as **customs and traditions** are concerned, festivities and how people celebrate certain traditions plays an important role in culture in general, but also in the conversations in TaCoCASE. Due to the different cultural backgrounds, the participants usually talk about customs or traditions from their own culture, explain cultural concepts related to festivities or traditions and compare it with the interlocutor's concept or take on these traditions. The participants negotiate meanings and exchange cultural concepts, which facilitates to a better (inter)cultural understanding and establishes common ground.

Customs and traditions are closely related with **cultural festivities or holidays**. Following the cultural identity theory, cultural practices and representations are a central aspect in the construction and negotiation of identity (see Hofstede 1993; Lüsebrink 2016; see also section 2.1.5.2.). Cultural festivals “are significant events that celebrate the cultural heritage, traditions, and values of a community” (Rutagand 2024: 10) and they mainly serve two functions: (1) as a “means to preserving cultural heritage” and (2) as a “platform for cultural expression and creativity” (Rutagand 2024: 10). Thus, celebrating cultural festivals contributes to appreciate cultural traditions and values and fosters identity (see Quinn 2013). Jack Santino, an academic folklorist and professor of popular culture in the U.S., who has published on rituals, celebrations and holidays, differentiates between festivals, which are “periods of communal celebration often centered around holidays” and holidays, which he describes as “day or period of days set apart from everyday life to celebrate important historical and religious people or events” or short: “holidays found on the calendar” (Santino 1996: 10). As far as holidays are concerned, he lists for example Thanksgiving and Christmas. Santino puts holidays in relation to pop culture and mass media and mentions several factors that are closely related to each other:

- Holidays and **books**: For example, on Christmas, when reading *A Christmas Carol* (Santino 1996: 19).

- Holidays and **religion**: Halloween, which is the evening before All Saints Day, or Christmas, which is a Christian holiday, celebrating the birth of Jesus Christmas) (Santino 1996: 57)
- Holidays and **music**: Certain songs are related to holidays, for example “Jingle bells” is usually played around Christmas (Santino 1996: 62)
- Holidays and **popular food**: Serving traditional food at specific holidays is very common, for example a turkey for Thanksgiving (Santino 1996: 65)
- Holidays and **politics** (Santino 1996: 70)

Besides these elements that are closely related to holidays, he also lists sending greeting cards as a tradition related to holidays. Sending greeting cards is very popular in England and the U.S. and it is usually done on family related holidays like Christmas or Easter (or Valentine’s Day) in order to send greetings to loved ones, which are usually family members and friends who are not around to celebrate the holidays with the ones who sent the card (Santino 1996: 93).

From an **EFL teaching** point of view, customs and traditions are a very important aspect of intercultural learning in the EFL classroom (Haß 2006: 146). Most syllabi of the federal states in Germany suggest to speak about festivities such as *Christmas*, *Halloween* and *Easter* already in early years of learning English (see Haß 2006: 146). Socio-cultural orientation knowledge should be established through discussing these festivities or holidays with regard to the history behind it, through music or dancing (Haß 2006: 146). In fact, most of the elements mentioned above that are connected to **holidays** can also be found in TaCoCASE, including music, food, religion and even sending greeting cards, as will be explored in the following.

In the data, customs and traditions are mentioned in eight conversations. The biggest celebrations or festivities that are mentioned are **Christmas, Thanksgiving, Halloween, St. Martin and Oktoberfest**. Talking about these festivities involves talking about the history behind them, some common practices as well as special food and drinks. The festivities will be described in the following and aspects the participants talked about will be explained very briefly. Most of the festivities will be subject to analysis following section (section 4.2.2.) with a special focus on negotiation meaning and discursive construction of meaning.

Christmas is one of the traditional festivities the participants talk about and that is mentioned in five conversations. This is actually not very surprising as the recording times were November and December and Christmas, as a Christian festivity, is celebrated in the U.S, Germany and the UK in December. Thus, the participants talk about a festivity that is close to them at the time of the recording.

In 05SB124BI14 (#1:15-6:00#), the participants talk about the Christmas season and the special mood they are in. They mention Christmas markets, which are usually found in Germany and the British speaker explains that there are Christmas markets in Birmingham as well, drawn on the German tradition. The participants discuss how they buy presents for their family and how difficult it is to choose good presents.

In 06BO18BI38, only the American speaker celebrates Christmas and so the British speaker wants to know more how the BO18 is celebrating it. BO18 explains that celebrating Christmas involves setting up a Christmas tree, having a Christmas dinner and BI38 joins the conversation by comparing how she perceives Christmas in the U.K (#1:00-1:30# and #3:45-5:00#). BI38 explains that she celebrates *Eid Mubarak*, which is usually celebrated after *Ramadan* among Muslims and she explains what the festivities involve (#1:30-2:00# and #5:00-6:30#).

In 06SB157BO11, the participants talk about Christmas as well and the participants tell how they celebrate it – since both are celebrating it – and what kind of special food they serve at Christmas (#4:20-7:40#). They compare their concept of Christmas – spending time with family, opening presents, decorating Christmas trees – with each other and discover similarities and differences. This example will be analyzed in detail later in this session (see example 14). Another interesting scene in this conversation is when BO11 brings up the German Christmas song “**Oh Tannenbaum**”, that she happens to know and that her father taught her when she was a child. She can even recall some of the lyrics and SB157 is very impressed that she knows the lyrics so well. This scene involves very much laughter and it is characterized by a very respectful and appreciative atmosphere (#23:45-25:00#).

In 07SB150BI32, a very special British tradition is mentioned in context with Christmas, namely writing Christmas cards (07SB150BI32 #14:07#).

In 06BO18BI38, the participants talk about **Thanksgiving** as BI38 asks BO18 how Thanksgiving was as it had just taken place. BI38 is interested in the history and the tradition behind Thanksgiving, which BO18 is happy to share in very detailed way (#2:00-3:45#). This example will be analyzed in greater detail later in this section (see example 18).

Halloween is mentioned in conversation 06SB157BO11 and in 07SB150BI32. In 06SB157BO11, Halloween is one of the participants' favorite holidays. SB157 says: "I really love Halloween because I'm really into the spooky sort of things". Halloween is traditionally celebrated in many Western countries on the evening before All Saints Day, which is referred to as All Hallow's Eve. In a pop cultural sense, it is associated with horror and spookiness. In the U.S., Halloween is celebrated with trick-or-treating and dressing up whereas in Germany, the celebration is usually only concerned with carving pumpkins and spookiness. Trick-or-treating however, has taken over from the U.S. and is also becoming a bigger issue in Germany.

In 07SB150BI32, Halloween is only very briefly mentioned as the start of a transition phase before Christmas, but the participants do not go very much in detail about how they celebrate it.

However, what is being discussed in 07SB150BI32 is **Saint Martin's Day**. Saint Martin's Day is usually celebrated on 11th November in some parts of Europe, particularly German-speaking regions, with processions led by a person representing Saint Martin on a horse that leads to a bonfire where everyone gathers and eats St. Martins Brezel. SB150 mentions Saint Martin's Day as it had just taken place and she assumes that her interlocutor does not know it. She explains the concept in great detail, including the procession, children making lanterns in day care, and the typical "Martinsbrezel" that people eat and thus, BI32 has an idea of the tradition and learned about it from her interlocutor (#10:50-13:15#). This example will also be analyzed in greater detail later in this section (example 16).

Oktoberfest, the Southern German traditional folk festivity, is discussed in 07SB156BO10 (#33:45-34:30#). The participants talk about typical or traditional food and when it comes to typical German food, SB156 mentions *Brezel* and *Weisswurst* and tells his interlocutor that it is very typical for Oktoberfest. BO10 knows the concept and knows what Oktoberfest is. She asks whether SB156 has a big Oktoberfest celebration in his city. SB156 explains that there is a celebration, but that the original celebration takes place in Munich. BO10 says that in America, they are trying to *copy Oktoberfest*, but she admits that it is mainly about drinking beer from her observation. The participants then talk about differences among German and American beer.

Some other customs, traditions, holidays or special events that are mentioned are mentioned in the following table:

TABLE 40: FURTHER ASPECTS OF CUSTOMS AND TRADITIONS IN TACOCASE

Independence Day, German Unity Day, Columbus Day	06SB157BO11; #9:00-17:20#
Veterans Day, Columbus Day, Labor Day, Valentine's Day	06SB157BO11; #41:30-44:30#
Mother's Day	07SB150BI32; #15:50-15:45#
Ramadan	06BO18B38; #1:30-2:00# and #5:00-6:30#
Black Friday (Sale)	08BO15BI36; #9:45-10:00#

The aspects that are depicted in table 40 were only discussed very briefly in the conversations and are mentioned here only for the sake of completeness. They will not be discussed any further as these holidays or special events did not occur very frequently in the conversations and it would be beyond the scope of this work to discuss every holiday or festivity that is mentioned in TaCoCASE in detail.

In sum, it is interesting to see that the participants talk about customs or traditions **from their own culture and from their interlocutor's culture** and cover mostly aspects that were found and established in research (see Santino 1996, Rutagand 2024). The majority of topics involves **festivities**, either (national) holidays, special days or religious holidays and include **music, religious aspects, and food**.

The participants typically start out by mentioning a **festivity** and asking whether the interlocutor knows it or not. Usually, they have at least an idea if the interlocutor could know it or not, which demonstrates that they have **sociocultural orientation knowledge** and knowledge about other cultures (see Byram 1997, 2021). When the participants talk about customs, traditions or festivities, they draw on their **skills of interpreting and relating** as well as **discovery and interaction** (see Byram 1997, 2021), which means they interpret other cultures, draw conclusions, compare it with what they know and what their concept is. The speakers **explain concepts** to each other, negotiate meaning of cultural concepts, which creates common ground and mutual understanding.

In contrast to intercultural communication analysis and critical incident theories (see Lüsebrink 2016), the participants in TaCoCASE usually do not have misunderstandings and engage in **successful communication**. This is due to the fact that the speakers are aware of the ELF intercultural setting of the conversations and they are aware that cultural concepts cannot be taken for granted in this kind of setting (see Cogo & Dewey 2012; see also section 3.1.2.2.4.). They also show openness and curiosity (see Byram 1997, 2021) as they usually want to know more about cultural concepts from their interlocutor's culture and ask for clarification. This draws on the attitude dimension in Byram's ICC model and demonstrates that the speakers act as intercultural speakers.

4.2.1.3. POPULAR CULTURE – TV SERIES AND MOVIES

TV shows, series and movies, sports and music are elements of popular culture that can be found in TaCoCASE and that are discussed among the participants. The participants, in general, have a hard time grasping the concept of popular culture, which one of the participants nicely put: "it is a popular culture, so I assume it could be anything" (BI36; 08BO15BI36, 22:28). This draws on the manifold definition of popular culture (see Storey 2006; see section 4.2.1.).

As far as **music** is concerned, the participants talk about different genres of music and what kind of music they like or dislike (03SB156BI24; #31:00-35:30#; or 08BO15BI36 #22:50-28:30#). As far as **sports** is concerned, in 06SB157BO11, the participants talk about sports on campus and especially football, which is quite a big aspect of identity for Universities in North America. The participants also talk about what kind of sports they like to do in their free-time or what kind of sports they are good at (06SB157BO11; #38:00-41:30#).

TV shows, series and movies will be subject to a more detailed analysis in the following. Before looking at how the participants talk about TV series and films in TaCoCASE, I will briefly touch upon previous research of TV series and on aspects of intercultural learning. TV series are one aspect of popular culture and they are as well transmitted through the mass media (Bednarek 2018: 3). Queen (2015) used the term “narrative mass media” to refer to film and TV series and in order to emphasize their narrative character. In terms of culture, TV series and shows are an “agent of socialisation” (Lippi-Green 2012: 101) and they affect our sociolinguistic environment (Coupland 2007: 185). The dialogue that is depicted (usually scripted) in TV series “constructs and reflects cultures and their ideologies” (Bednarek 2018: 3). There has been conducted some research on the language in the media, in particular in TV series and films (see Androutsopoulos 2012, Queen 2015, Bednarek 2018).

Valentin Werner (2018), who is editor of the book *The Language of Popular Culture*, states that the **language in popular culture** had for a long time been ignored in English linguistics and “underexploited in (second) language instruction”, and he thus sees great “didactic potential” for investigating the language of popular culture in EFL teaching (Werner 2018: 3). In his edited volume, authors for example look at music and lyrics, linguistic representation of gender roles as well as how identities and cultural discourse are represented in pop culture with a focus on language. The language in popular culture and in particular in film and TV series has been subject to analyses in some studies (see Androutsopoulos 2012, Queen 2015, Bednarek 2018). In terms of content of TV series and films, it should be referred to the characteristics of TV series, having “two information tracks” (Chatman 1999:

318; emphasis in original). These two tracks can be described as follows (Allrath et al. 2005; adapted from Chatman 1999: 135):

- **Auditory channel:** noise, voice, music, earshot, commentative
- **Visual channel:** actor (appearance, performance), location, prop, cinematography (camera, colour, lighting, mise-en-scène), editing

These aspects already give an impression what factors could be subject to an analysis of TV series other than language.

TV series in EFL teaching research have mostly been analyzed and used with regard to establishing communicative competence. There are studies on the vocabulary development of watching TV series (Shabani & Zaanussi 2015), establishing speaking and listening skills through TV-series-based instruction (Kung 2013) or using TV series to develop pragmatic competence (Ángel Villegas et al. 2022), just to mention only a few. There is a strong focus on communicative competence and only a few studies look at how TV series can support intercultural competence. One example is a study by Larrea Espinar and Raigón Rodríguez (2019), which looks at culture in sitcoms like *The Big Bang Theory*, *Friends* and *Modern Family*, among others.

In terms of **intercultural learning**, EFL teaching methodologies suggests working with literary texts (drama, novels), short stories, films, comics, advertisements and cards (Haß 2006: 147; see also Decke-Cornill et al. 2015). **Films** seem to be established in intercultural learning, even though they were seen controversial at the beginning (see Decke-Cornill et al. 2015: 246). As far as movies or films are concerned, they serve four main functions in the EFL classroom (see Decke-Cornill et al. 2015: 246; my translation):

- (Potentially educational) film experience
- Language acquisition and learning, as well as knowledge of the country
- Film analysis skills
- Identity creation

Thus, these elements draw on communicative as well as intercultural (communicative) competences and additionally, to media competence. In a similar manner, TV series could as well be used for intercultural learning in the EFL classroom.

As far as **content** of the discussion of texts (literary as well as media texts, including film and TV series) are concerned, Armstrong (2005), who focused on short stories in intercultural learning, came up with three elements that should be analyzed: language, culture and metacognition. He lists several factors that help investigating these elements (Armstrong 2015: 13):

- setting
- characterization
- plot/structure
- theme
- context
- narrator's point of view

In discussing **short stories** with the learners, the famous “5W” could structure the analysis: “what, who, where, when, why, how” (Armstrong 2015: 14). Discussing and analyzing short stories in the EFL classroom raises cultural awareness and allows cultural comparisons (Armstrong 2015: 16), which is true as well for any other kind of text, including **TV series and films**. TV series are close to the learners’ real-life experience and it could be very motivating for learners to discuss these topics from their leisure time also in the classroom. TV series and movies therefore have great potential and should be used in the EFL classroom (see section 5.3.).

In fact, in TaCoCASE, the participants talk about TV series and mention some of the factors, such as **setting, characters and plot or theme**. There are also new factors, for example the medium where people watch these series. It is important to note at this point that not TV series themselves are investigated in the following, but how the participants in TaCoCASE talked about TV series. This means, the conversations contain not only factual knowledge about the setting or characters of the TV series, but also about personal attitudes and preferences towards the series. This shift of perspective and including attitudes, which is one of Byram’s *savoirs* of ICC (Byram 1997, 2021; see also section

2.1.5.2.3.) that is often neglected in intercultural learning (see König et al. 2022: 8). In using the conversations about (pop) cultural aspects of TaCoCASE, a more discourse-oriented and intercultural learning in the EFL classroom is suggested, which is demanded by many prominent intercultural learning and EFL teaching scholars in Germany (see Hallet 2008; see also König et al. 2022: 8).

As far as **TV shows, series and movies** are concerned in TaCoCASE, there are several examples the participants discuss. Interestingly, it is not only the show, series or movies themselves that the participants talk about, it is also a meta-discussion about **reasons for watching** or certain aspects from these movies, series or shows, such as **celebrities or actors** involved.

The TV series, shows and movies that are mentioned and discussed are depicted in the following table:

TABLE 41: TV SHOWS, SERIES AND MOVIES IN TACOCASE

03SB149BO03	Time in conversation
Saturday night live, King of Queens	#07:30-08:45#
05SB124BI14	
Deadpool, Doctor Strange	#13:15-17:00#
Titanic	#52:40-55:00#
Matilda	#56:00-59:00#
07SB150BI32	
Disney movies (and Disney characters)	#04:20-05:40#
Sherlock, Desperate housewives, Big Bang Theory, the Goldbergs, Gossip Girl, Gilmore Girls, Pretty Little Liars	#37:50-41:00#
08BO15BI36	
Hannibal, Bones, The Walking Dead	#01:00-03:45#
Indian drama TV show (triangle)	#03:45-04:40#
Friends (favorite character)	#04:40-05:50#
Hannibal	#06:25-07:20#
Inception	#07:20-07:45#
Stranger Things	#08:15-09:20#

Reality TV (parents, cooking shows)	#10:00-11:30#
Sherlock, Doctor Strange	#14:30-15:45#
Nikita	#16:50-17:50#
Marvel movies	#17:50-18:30#
Anime	#18:30-21:00#; #31:50-34:25#; #42:45-46:00#
Disney films	#21:00-22:00#
Christmas movies: Elf, Grinch	#39:50-41:30#
Marvel (poster in the background)	#48:00-51:00#
08BO18BI41	
Simpsons, Family Guy, Firefly, Game of Thrones, Frasier, The West Wing, How I met your mother	#02:25-08:30#
Dr. Who	#08:30-13:10#
Horror movies	#13:10-13:25#
90s series and movies, Deadpool, superman in the 70s versus nowadays	#14:45-21:20#

This list consists of mostly **American and British TV series** that are available on streaming platforms like Netflix. It seems that Western and in particular British and American TV series and movies are predominantly mentioned by the participants. This is probably due to the fact that the TV series and movies are in English and are thus easier to understand and they are very widespread. It could also be a sign of intercultural competence that the participants chose TV series that they think the interlocutors could know as well to establish common ground and talk about something both can relate to in order to prevent a communication breakdown or topic shift.

The participants address the **following topics** when talking about these **TV shows, series and movies**:

- Celebrities: actors and actresses (*Kate Winslet, Benedict Cumberbatch, ...*)
- Different genres
- Differences across countries (*what's popular in..*),
- Medium (*Youtube, TV, Netflix, ...*)
- Reasons for watching: want to relax, re-watching series

- Recommendations: *“you should watch it”*
- Stereotypes portrayed in TV series
- Community / “coolness”

It is interesting to see that most of the elements that Armstrong (2015) mentioned – when discussing short stories in the EFL classroom – can also be found in the conversations in TaCoCASE. The results indicate that the content of the conversations in TaCoCASE is quite similar to what has been found in research (Armstrong 2015) even though the medium is different. The participants in TaCoCASE talk about TV series and movies with regard to characters, in this case also the actors that play the characters in the TV series and they talk about different genres of TV series and movies (for example horror movies, romance). Some aspects that they discuss are new – due the medium – and include cultural or affective issues.

The **medium** plays an important role in talking about TV series as the medium provides access to TV series. It is not only TV, it is also streaming platforms like *Youtube* or *Netflix* that the participants mention that they use to watch TV series or movies. This is a new aspect that is usually not discussed with traditional media or texts, like books for example.

As far as (inter)cultural issues are concerned, the participants talk about **stereotypes** and how they are portrayed in the TV series as well as differences across countries and what series are popular in certain countries. This usually difference in popularity across countries usually has to do with the origin of the series, for example German series tend to be very popular in Germany and not known in other countries because of the (German) language – if they are not dubbed. American and British series are more widespread and known as the language issue does not exist and people from various countries watch these series in English even though it is not their mother tongue.

Another cultural issue is more an identity creating issue like a sense of belonging to a certain community (of practice) or social group that watches certain TV series. One of the speakers said: *“I’m not cool, I don’t watch a lot of the newest things, I haven’t even seen an episode of (..) Game of Thrones”* (08BO18BI41;#06:30#). She has a sense of not belonging to a certain social group of people that could

be described as the community of *Game of Thrones* or in general the community of people that watch the latest TV series. She expresses this feeling of being an outsider and not belonging to this social group or community of practice by saying that she is not “cool”, which means that she identifies pop culture and watching the latest series as being cool (see for example Fellner et al. 2014 on coolness and affective encounters with American popular culture). TV series and movies are thus identity creating issues and people not only **identify with certain TV series** and distance themselves from others, but also from social groups that watch these series and that are communities of these series.

The last two elements the participants talked about are of affective nature and show **attitudes towards TV series and movies**. This is also to some extent a cultural issue, but it is also a very personal issue. The participants for example name **reasons for watching TV series** or movies and they **recommend** them to their interlocutors. The participants talk about personal preferences, what they like or dislike and take stance with regard to certain TV series and movies. In talking about these personal preferences, the participants show affective behavior and talk about their personal attitudes.

This is only a very brief analysis of the elements that are mentioned in TaCoCASE and the discussion and analysis only represents the start of a much broader debate and research that should be continued in the future. It is clear from this analysis that TV series and movies themselves but also conversations about TV series and movies are very well suited for intercultural learning. TV series and movies contribute to cultural understanding and awareness, are close to the learners’ real world and initiate discussions and issues of identity and belonging to social groups or communities.

In particular, talking about TV series and movies corresponds to two main principles that contribute to the shift in intercultural learning towards establishing ICC in a modern EFL classroom: (1) from culture-as-a-nation to interculturality and (2) from being to doing culture (Kramsch & Hua 2016). This transcultural as well as discursive perspective that talking about TV series and movies has, is very promising and new and should be investigated in further research. The examples from TaCoCASE in which the participants talked about TV series and movies provide a rich source for intercultural

learning as not only factual knowledge about countries or stereotypes are discussed, but also attitudes towards series and how people like or dislike or identify with certain TV series and movies. Even though the latter is one central aspect of Byram's ICC model, attitudes are usually not taken into account in establishing ICC in the classroom (König et al. 2022: 8). The following section looks closer at how the participants talk about cultural aspects.

4.2.2. TALKING ABOUT CULTURAL ASPECTS

RQ6: How do the participants talk about cultural aspects?
--

Talking about cultural aspects refers to a new model of ICC teaching or intercultural learning, described as **discursive intercultural learning** (Altmeyer 2017: 12; see also Hallet 2008; Kramsch & Hua 2016). While textbooks, literary texts as well as movies are usually used in intercultural learning, examples of interaction about cultural aspects are rare (König 2022: 20). The examples from TaCoCASE represent interaction of intercultural encounters and are therefore a great basis for understanding how cultural concepts are negotiated. This process of negotiation of meaning in intercultural encounters involves **mediation** and exchanging **(new) cultural concepts**.

This kind of mediation can be referred to as “cultural mediation” (Gohard-Radenkovic et al. 2004), which is also the basis for the concept of mediation in the CEFR when learners should be able to explain new concepts. **Cultural mediation** has not been discussed very widely in research (see Gohard-Radenkovic et al. 2004), even though it is very similar to the concept of negotiating meaning (Hua 2016). In cultural mediation, clarification of meaning is one of the key strategies that contribute to successful intercultural communication (Gohard-Radenkovic et al. 2004: 48).

Communication strategies to negotiate meaning have been researched both from a learner language research perspective as well as from EFL research perspective (see Long 1996; see also section 5.2.). Communication strategies are defined as “a mutual attempt of two interlocutors to agree on a

meaning in situations where requisite meaning structures do not seem to be shared” (Tarone 1981: 288). Communication strategies that have been identified in learner language research, usually **among NNS**, are paraphrasing, reformulation, requesting help or clarification, approximation, transfer, avoidance, discourse strategies (see for example Long 1996, Tarone 1981, Faerch & Kasper 1984, Dörnyei 1995, Celce-Murcia et al. 1995, Dörnyei & Kormos 1998). Communication strategies in **ELF research** focused on clarifying meaning, pre-empting and resolving misunderstandings, discourse markers and laughter (see Wagner & Firth 1997; Kaur 2010; Mauranen 2006, Pietikäinen 2017; Pitzl 2010; House 2013; Buysse 2012; Meierkord 2000). As far as ELF conversations and especially corpus-based analyses are concerned, Brunner (2021) investigated communication strategies in ViMELF, the sister corpus of TaCoCASE, including innovative language use, code-switching, definitions and non-verbal behavior. Antonello (2022), who also looked at communication strategies in ViMELF (and VOICE), investigated backchannels, multilingual/plurilingual resources, lexical anticipation, lexical suggestion and correction, reformulation, repetition and spelling.

In my analysis, I take an interactional perspective on communication strategies and focus on how **meaning is constructed discursively and as a collaborative process in interaction** (see Gumperz 1982; Hua 2015; Baker 2015). Therefore, ELF communication strategies and in particular strategies to clarify meaning (see Wagner & Firth 1997; Kaur 2010) are subject to analysis in this section. Other strategies, such as repetition, spelling, reformulation or code-switching (see Long 1996, Antonello 2022; Brunner 2021) will also be subject to analysis when it comes to discourse structure and how the participants explain (new) concepts (see section 5.2.2.).

In TaCoCASE, the participants talked about cultural concepts and **explained (new) cultural concepts** from all areas mentioned above, including food, customs and traditions and popular culture, such as TV series or movies. Thus, a cultural concept in this case could be a food item, a special festivity or tradition or a TV series or a movie, for example. The only difference is that food and customs and traditions are usually culturally bound to countries or to the speakers’ culture, whereas TV series and movies are more of a global nature and can be regarded as transnational aspects. This means that the

steps “concept from own culture” and “concept from interlocutor’s culture” can be ignored for this part. The participants need to clarify the meaning of the (new) concept and apply different strategies to do so. The strategies they used generally base on the communication strategies mentioned above, like clarifying meaning, and they are summarized in figure 51.

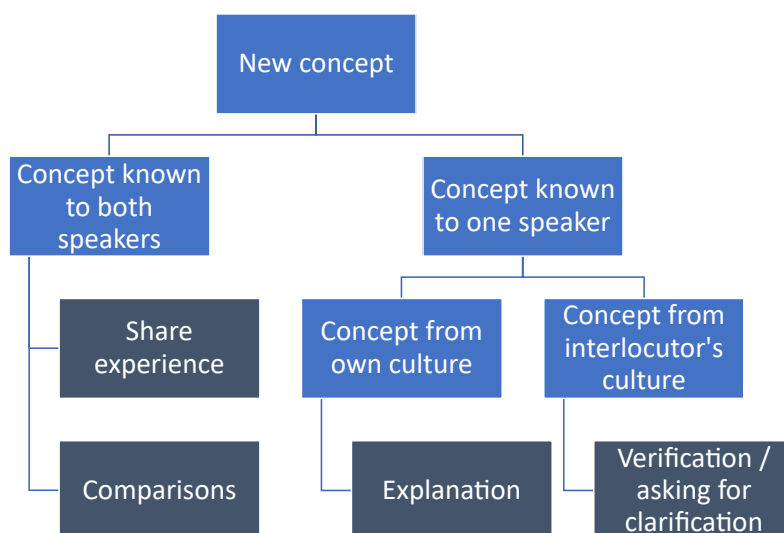


FIGURE 51: STRATEGIES FOR MEDIATING CULTURAL CONCEPTS

In general, I found two cases that should be mentioned here: (1) the concept is known to both speakers or (2) the concept is only known to one speaker, which means that the concept is unknown to one of the participants and one of the participants has a lack of knowledge. This lack of knowledge needs to be filled, which is usually done with the help of communication strategies like clarifying meaning or other mediation strategies (see section 5.2.).

If the **concept is known to both speakers**, there is no information gap between the speakers and the speakers usually start **sharing their experience** with the respective concept, be it a festivity, food or a TV show. This concept could be either from their own culture, from the interlocutor’s culture or a “third” culture they both do not identify with, but know about the concept. If the concept is a transnational phenomenon, this cultural differentiation “own versus third culture” can be ignored. Especially when it comes to customs and traditions or food, the participants usually **compare** their

respective concepts with each other, talk about differences and similarities, which contributes to the exchange of concepts. This is mediation and intercultural communication at its best as it contributes to mutual understanding and creates common ground between the participants.

Example 14 is an excerpt of the **comparison** that the speakers make between their way of celebrating Christmas.

(14) Concept known to both speakers: Comparison – Christmas (06SB157B011; #04:20#)

- SB157: =a:h okay.
it's- it's Christmas soon {claps hands together} so uhm,
- BO11: YA:Y.
- SB157: so I-I guess you celebrate it from that reaction [((chuckles))]
- BO11: [YEAH ((LAUGHS)) yeah].
well I guess yeah [((lauhging))].
- SB157: [do you-],
- BO11: [well it's] Christmas really no (honestly) [((laughing))].
- SB157: [((laughs)) I "do" know a] couple of people who really just detest Christmas. #04:30:00#
and they don't- they don't celebrate it.
[well I can]-
- BO11: [oh really]?
- SB157: yeah it's yeah that's somehow how it is.
I mean I think it's sad because it's a **really nice,**
holiday to be with your family and friends,
uhm how do you usually celebrate it like uhm,
are you together with your family on that day?
- BO11: I'm "usually" uhm you know,
put up the Christmas tree,
and I'm not really religious (personally) but uhm it's kind of,
in America at least It think it's like a non-
... (denominational) holiday and we're like, #00:05:00#
pretty much everybody celebrates ((chuckling)) Christmas,
even if they're not religious so,
uhm ... t yeah we **put up the Christmas tree do presents** and,
it's always fun **to be around kids,**
when they're **opening presents** on Christmas day [((laughs))],
- SB157: [((chuckles))].
- BO11: so that's my favorite part [but],
- SB157: [aw:h ()],
- BO11: [yeah how do] you celebrate?
- SB157: u:hm ... {looks in the air}
it's really basic I'd say so uhm,
we **celebrate on the twenty-/firth/- er twenty-"fourth"** already,
- BO11: .. [oh .h].
- SB157: [so] one day **tha- on the night before you do** ((heh)), #00:05:30#
- BO11: [ah].
- SB157: [<and>] we usually **put up the Christmas** tree together as well,
it's mostly my brother and I who have to do the work,
in that ((hehe)) [apartment ((laughing))].
- BO11: [((laughing))].
- SB157: <and> then we go to our grandmother's,
<and> we basically just spend the evening there,

have a **large dinner sing**,
 sometimes we **sing Christmas songs** or my brother plays on his violin,
 and then,
 BO11: [oh nice].
 SB157: [it's time] for **presents** of course ((ehh)).
 BO11: ... aw that's fun ((laughs)).
 SB157: [(it is really)]-
 BO11: [so you] open presents on Christmas Eve.
 or like-
 SB157: exactly. {nods}

As this example demonstrates, both participants celebrate Christmas and express cultural expertise about the holiday. There are similarities and differences of how the participants celebrate Christmas. For both participants, Christmas is a family related holiday (see Santino 1996) that involves celebrating with the family and follow family traditions: putting up a Christmas tree and preparing presents. However, the participants also discover differences when it comes to opening the presents: In Germany, the presents are usually opened on the 24th December and in the U.S (as well as in the UK), the presents are opened in the morning on Christmas Day (25th December). In fact, the German speaker shows cultural expertise about the American culture as she brings up the difference (*we celebrate on the twenty-/firth/- er twenty-"fourth" already*) and the American speaker shows a lack of knowledge by saying “oh”, which indicates that she is surprised that there is a difference regarding the time when to open the presents. The American speaker expands her knowledge on the German tradition and she signals understanding and confirmation in drawing the conclusion and repeating that Germans open the presents on Christmas Eve. Both participants thus **compare** their concept, both assuming the role of cultural experts and explaining their cultural concept of Christmas in their home country or their family traditions to their interlocutor.

In example 15, the participants talk about a concept that they both know but that is from a culture they both do not identify with. The concept the participants talk about is *ramen*, a Japanese noodle dish.

(15) Concept known to both speakers: Share experience – Ramen (07SB157BO11; #35:00#)

BO11: [...]I would like to like have,
 shoyu ramen ((Japanese (0.5)) and sushi ((Japanese ((0.5))),
 [and yeah] ((hehe))
SB157: **[o:h] (miso) ramen** ((Japanese (0.5))), ((chuckles))
BO11: a:h yes ((laughs))
SB157: so good,
BO11: there are there are a couple of places here,
 that make **like authentic style ramen** ((Japanese (0.5))),
 but **I don't know how authentic it is**, {shakes head}
 >I don't know<,
 I'll have to check it out.
SB157: () half authentic {measures length of something with fingers} ((chuckles))
BO11: oh too good (out there) ((LAUGHS))
 .. yeah I really want to like #00:35:31-6#
 go there for a year and like get better at Japanese,
 (I wish I could do that),
SB157: mhm, ((hehe))
 [(ganbatte)((Japanese (0.5))), {lifts thumbs up in the air}
BO11: [yeah]
 °O:H° (oh man)((Japanese (0.2)) yes ((laughing)),
 thank you .. I appreciate that ((laughing)).
 (aligato mas) ((Japanese (0.5)) ((laughing)),
SB157: (namaste)((Japanese (0.2)) ((laughing)) {makes waving gesture}

BO11 introduces the concept by stating that she would like to have *shoyu ramen*. SB157 is excited when she hears the word and uses an interjection (oh) to signal her surprise and excitement. She immediately thinks of *ramen* and she expresses that she likes ramen (*so good*). BO11 talks about places that make authentic style ramen in her city, but she is not sure how authentic it really is, because she has not been to Japan to taste the original ramen. She is using hedges and vague language and is thus not claiming expert status. They both share their experience of tasting ramen. They even end up speaking Japanese to each other, which is an intentional code-switch in order to show their mutual respect and involvement with Japanese culture and language. It could be argued that they are creating a third space here, where they both identify as learners of Japanese and show their attraction to the Japanese culture. This shows again how close language and culture are intertwined and that two people understand each other and create common ground by creating a common space where they talk about a dish from a culture both know and like and using a language from a culture and country they both like.

If the **concept is only known to one speaker**, there is an information gap between the participants in the conversation that needs to be overcome. The speakers usually use (cultural) mediation, as explained above, to clarify and negotiate meaning of cultural concepts (Gohard-Radenkovic et al. 2004). Sometimes, the speaker who does not know the concept has heard of it or has at least an idea of it, but sometimes the concept is completely unknown to them. In general, there are two cases for when the speaker brings up a new concept: (1) Either the concept is from his or her own culture or (2) the concept is from the interlocutor's culture – except for transnational concepts.

If the **(new) concept is from the speaker's own culture** and the interlocutor does not know it, he/she usually explains the concept so that the interlocutor understands it and has an idea what it is. The speaker has expert knowledge and claims cultural expert status. In explaining the concept, the speaker shares his or her expert knowledge with the interlocutor and the interlocutor thus expands his or her cultural knowledge. In the following example, the German speaker introduces the concept Saint Martin's Day to her British interlocutor.

(16) Concept known to one speaker / own culture: Explanation – Sankt Martin (07SB150BI32; #11:00-13:00#)

- SB150: [...] we've got something else as well,
<in November> .. I don't know if you know that,
it's called **Saint Martin's Day** ...?
- BI32: u:hm no [I don't],
- SB150: [it] doesn't ring a bell .. no,
I thought that,
because it's kind of a typical German thing,
I would say .h,
- BI32: {nodding} yeah,
- SB150: it's more like u:hm ... #00:11:00#
it's for children,
es- especially for CHILDREN,
but also adults are going to ENJOY it .. usually,
it's more like .. it's one day.
I think it's usually **around the eleventh of November,**
a:nd u:hm you are **going to church** then,
- BI32: {nodding},
- SB150: and afterwards there's a **procession,**
kind of a procession,
if you want it or <parade>,
I would either .. uh >would<,
- BI32: mhm,
- SB150: .. say process- ,procession,
a:nd u:hm all the people are gathering,
in front of the CHURCH,

when it's already dark,
a:nd u:hm children prepare u:hm-, #00:11:30#
I think in day care,
BI32: {nodding},
SB150: or wherever they went to .. u:hm during the day.h,
u:hm but they u:hm,
basically prepare kind of **lanterns**,
BI32: {nodding},
SB150: could be anything,
could be <like a vulpine or a hedgehog>,
BI32: {nodding},
SB150: leaves .. whatever .h,
a:nd u:hm in these lanterns, {moves hands from inside to outside imitating opening a lantern}
BI32: {nodding}
SB150: there kind of a glow stick,
BI32: mhm,
SB150: so they put it in there,
and when it's then uh:=,
=when it's DARK u:hm,
BI32: {nodding},
SB150: they put it <o:n>,
a:nd then everyone is u:hm=,
=**starting to <SING>**,
special songs and then we are, #00:12:00#
actually walking,
and it's- is-h [It's basically just],
BI32: [uh: mhm],
SB150: u:hm being <together a:nd singing>,
a:nd uh- for the children having lots of fun,
and joy .. being cheerful,
a:nd u:hm .. it's kind of a short procession,
BI32: {nodding},
SB150: AT LEAST in the village where I'm coming from,
.h and u:hm we gather on a town square then,
there's a **huge fire**, {moves arms around}
BI32: yeah,
SB150: u:hm .. so we enjoy uh: <the fire and the warmth>,
and it has also something special that we EAT,
at these- u:hm around these days.. , #00:12:30#
or at this day .. H uh: doesn't matter,
ant that's- it's like u:hm ..,
you could imagine a <gingerbread puppy>,
do you say gingerbread puppy?
or gingerbread [man]?
BI32: [uh:] gingerbread man ((laughs)).
SB150: yeah .. basically ..,
but it's u:hm made out of yeast,
BI32: yeah,
SB150: it looks lik- .. a bit like this,
it's made out of u:hm yeast pastry,
a:nd u:hm it's with sugar icing,
and children <really> >really< love it,
it's like- it's called Brezel ((German (1.0))),
so brezel ..,
BI32: o:kay .. yeah,
SB150: **Martinsbrezel** ((language German 1.0)),
BI32: {nodding},
SB150: uh: yeah and they really love it, #00:13:00#

The German speaker introduces the concept Saint Martin's Day, translating the German holiday into English, to check whether her British interlocutor knows about it as well. The British speaker does not know it (*doesn't ring a bell*) and SB150 is assuming cultural expertise and is explaining the concept in a longer stretch of utterance what Saint Martin's Day is about to her interlocutor. The German speaker is also showing socio-cultural knowledge when bringing up the concept as she assumes that her interlocutor does not know the concept as it is a German festivity (*I thought that, because it's kind of a typical German thing*). SB150 is identifying herself with the German culture by using the word *we*, indicating that she is speaking from experience and also celebrates Saint Martin's Day (*we gather on a town square then*). She starts off by explaining when the holiday is celebrated (*around the eleventh of November*) and then starts describing what people do on that day: people go to church and there is a procession with singing involved, children prepare lanterns and the procession ends at a bonfire. There is also a special kind of pretzel that is served, which is called *Martinsbrezel*. It is interesting to see that SB150 is explaining most of the time while BI32 is not speaking very much, in fact BI32 is only backchannelling to signal understanding (*yeah, nodding, mhm*) and to indicate interest (*okay yeah*). SB150 involves BI32 from time to time in the explanation when asking about a correct word, either because she does not know the word and asks for the word or because she wants to distract from her cultural expert status and acknowledges the linguistic expert status of her interlocutor. Other than that, she holds the floor and explains the concept with many pieces of information in order to be very precise about the concept even though it is clear that she has a hard time to clarify meaning as she uses vague language and hedges (*kind of, something special*), probably because of lack of vocabulary and innovative language use (which is common for ELF speakers when negotiating meaning; see Brunner 2021). The German speaker shows cultural expertise through evaluative language use (*children <really> >really< love it*). In this example, we can also see how a festivity is described through practices, food and music (see Santino 1996; Rutagand 2024).

The *Sauerkraut* example (example 19) is another good example, in which one speaker is explaining a concept from his or her own culture. This example will be analyzed later in this section with special regard to discourse strategies.

If one of the speakers introduces a **(new) concept that is from the interlocutor's culture**, the speaker has usually heard of the concept but is not sure about the details of the concept and so, the speaker either tries to **verify** the concept he or she has in mind or **asks** the interlocutor – as the cultural expert – **for clarification**.

(17) Concept from interlocutor's culture: Verification / Asking for clarification – Christmas
(06BI38BO18; #00:30#)

BI38: u:hm,
...so, #00:00:30-6#
do you celebrate CHRISTmas?

BO18: .t uh ja,
I do.
uh do you?

BI38: **NO: but it seems j- it seems like a really nice holiday though.**
[it's not] the-

BO18: [I'm not]
I didn't enjoy it.
immensely actually.
((clears throat))

BI38: **so [what do you]**
.. in Christmas?

BO18: [I don't know]
well uhm,
... set up a **christmas tree** (noise in background)
.t **decorate it with the family,**
[uhm],

BI38: [oh that's] nice

BO18: .t usually there is a **big dinner,**
either at Christmas [eve or-or]

BI38: **[IT'S j- IS there ACTually a] real tree though?** ((very loud and distorted))

BO18: what's that?

BI38: be- is it- IS it a real tree,
because I think in,
.. UK households, #00:01:00-6#
.. it's fake.
so.

BO18: oh.

BI38: don't know,
[it's j-]

BO18: [oh]

BI38: yeah

BO18: and it depends
uhm .. f-

BI38: okay.

BO18: we've had real trees,

we:-
 uhm .. my parents are divorced,
 so I have a tree in each house.
 uhm .. .t.
 BI38: >mh<
 BO18: and u:h
 >it kind of depends<,
 some times my dad will like to do a real tree,
 uhm but,
 other times,
 MOST of the time we do a fake tree.
 and at my mom's house we PRETTY much always do a fake tree,
 I think we only had a real tree once or twice. .h
 BI38: .. okay.
 BO18: **what sort of holidays do you celebrate?**
 BI38: uhm
we celebrate E:ID. #00:01:29-9#

In this example, it is interesting to see that BI38 brings up the concept Christmas and asks her interlocutor if he celebrates Christmas. BI38 signals general knowledge about the concept but she says that she does not celebrate Christmas, which makes her a cultural non-expert in this case. She seems interested in how BO18 celebrates it and asks the cultural expert about the festivity to verify her concept and ask for clarification. She asks what BO18 usually does on Christmas and BO18 explains that they set up a Christmas tree and have a dinner. BI38 interrupts BO18 with a question about the Christmas tree, to signal active listenership and to ask for clarification. She asks if it is real tree or not as she observed in the UK that most of the people use fake trees. She thus draws on prior knowledge and compares the information she has to the expert information that her interlocutor can provide. She thus signals co-expertise to some extent and tries to verify her observation and compare British Christmas trees to Christmas trees in the United States. She thus asks her interlocutor for clarification, who is the expert for her in terms of Christmas tradition at this point. He uses hedges and vague language (*it kind of depends, most of the time, pretty much*) and hesitation markers to downgrade his expert status and to signal that they are both at equal levels. It could also be that he wants to emphasize that he is describing a personal preference, which could be different from the experience his interlocutor had. He acknowledges co-expert status for his interlocutor based on the experience and observations she had and he is not claiming expertise on this particular item of the festivity that he is claiming general cultural expertise of. This example shows also very well how BI38 is trying to

complete the image of the concept that she has in mind by asking about aspects of the concept and verify the concept. In asking for clarification and comparing the image in her mind to the concept that the interlocutor explains, she shows very much interest in her interlocutor's culture and his Christmas practices and signals cultural awareness and sensitivity. This is echoed by the American speaker after the explanation of what BO18 does at Christmas, when BO18 asks BI38 what sort of holidays she celebrates in order to signal interest. Throughout the conversation, both participants thus show great openness and curiosity in the interlocutor's culture as well as cultural sensitivity and awareness.

Another excerpt from the same conversation that could be used here is the **Thanksgiving** example (18), where the British speaker asks the American speaker about Thanksgiving as she does not know about it (06BI38BO18 #02:00 – 03:32#).

(18) Concept from interlocutor's culture – Thanksgiving (06BI38BO18; #02:00#)

- BI38: O:H you had **THANKSgiving**,
what was that like,
'cause I don't know what Thanksgiving is? #00:02:00-8#
- BO18: OH okay,
 well so the story behind Thanksgiving is u:hm
t ... way back whe - do you know the story or am I boring you with that
- BI38: >No no no< I don't know,
all I know is there is a turkey
- BO18: [oh]
- BI38: [and] that is it
- BO18: yeah that's - that's pretty common,
 so uhmt .. so yes the-the story is way back when,
 uhm the pilgrims came,
 uh from England,
 due to religious persecution,
 .t .. uhm .t and when they got here .. they uhm,
 ... they were kind of running out of food,
 so they talked to the Native Americans, #00:02:30-4#
 and the Native Americans agreed to share some of their,
 .h ... uhm harvest
 and >saved them<,
 and uh .h,
 their idea is that they had uh one big turkey,
 .t uh to feed everybody ((clearing throat))
 .t and that's how they survived the WINter.
 so u:hm.
 ... t. the last Thursday of November,
 we uhm,
 .. or third Thursday of November actually,
 excuse me,
 uhm we get together and uh,

.t .. we have turkey cause they're supposed to have eaten turkey.
 uhm,
 BI38: oh okay
 BO18: and then, #00:03:00-8#
 all bunch of stuff.
 uhm,
 .t stuffing cor- I mean it depends on where you are,
 I know plenty of people who actually eat LAMB,
 uhm several others who make HAM
 .t ... u:hm .. that sort of thing so,
 .t yeah uh it's a pre- it's a-
 .. one of the holidays everybody is supposed to get together.
 you know your whole family is supposed to come into TOWN,
 that sort of thing.
 BI38: is it bigger than CHRISTmas though?
 is it kind of the same,
 BO18: uhm I would say it is .. it's on part,
 probably slightly <LESS> important
 in most people's minds than Christmas
 BI38: oh_okay. #00:03:32-1#
 BO18: yeah
 uh it's- still is pretty important, {sound in background}
 you still- you normally will travel,
 to see your family,
 but ... not AS important as Christmas.

BO18 and BI38 are very nicely negotiating what they would like to talk about as BO18 is about to start talking about the history behind **Thanksgiving**, but then asks BI38 if he should talk about history or if it is too boring for BI38. BI38 admits that she only knows about the turkey, thus signaling non-expert status and signaling that she wants to know more about the history and other details. She signals lack of knowledge and thus prompts the explanation by the cultural expert. BI38 shows great openness and curiosity about BO18's culture and practices. BO18 is happy to explain the history of Thanksgiving in a very detailed and nuanced way. After that, BO18 explains when Thanksgiving is celebrated (*on third Thursday in November*) and he then explains what they usually eat at Thanksgiving as the Thanksgiving dinner with the family is an important part of the festivity. We can see here that food plays an important role in celebrating festivities or holidays (see Santino 1996) and that when the festivity takes place and what people is an integral part of explaining a cultural concept or festivity (Santino 1996; Rutagand 2024).

It is interesting to see that both examples involve **verification or asking for clarification** from the non-expert as well as explanation from the expert. The cultural expert explains the concept they are asked

about and there is a certain pattern that this explanation follows. As the Thanksgiving example (18) and the Saint Martin's Day example (16) show, there are **four elements** that are usually mentioned when **explaining a (new) concept**:

- History (what is celebrated)
- Description of practices (what people do)
- Food (what people typically eat)
- and the date of the festivity (when it is celebrated).

These elements are in line with elements that were found in established research about festivities and holidays, such as history, practices and rituals as well as food and the date or the season of the holiday or festivity (Santino 1996; Rutgand 2024). Sometimes the participants in TaCoCASE also provided a personal evaluation, like for example what aspects of the festivity are perceived in which way by children or adults. This is a new kind of information that had not been addressed in previous research very much. This is probably due to the intercultural setting TaCoCASE is set in and also the discursive construction of meaning that the participants feel to add a personal note or evaluation to the explanation in order to emphasize their expert status and in order to set the festivity in a positive light, which makes it appealing for the interlocutor. The participants are in a way serving as cultural ambassadors in explaining the cultural concepts and "selling" them to their interlocutors. This is an aspect that can be found in TaCoCASE and that should be further discussed in future research.

As we have seen from the examples, there are **certain patterns** of discourse that evolve from the examples. The discourse structure that the participants used when explaining cultural concepts will be investigated further in the following.

Discourse structure of explaining (new) concepts

The examples suggest that when people exchange and mediate cultural concepts, the **discourse structure** follows the same **basic pattern**.

The following table (table 42) displays the discourse structure that had been identified in TaCoCASE when participants talked about cultural aspects and explained (new) cultural concepts. The left column displays the identified discourse element and in the second column, the function of the respective discourse element is explained. An example of expressions the participants used in combination with the discourse elements is depicted in column 3.

TABLE 42: DISCOURSE STRUCTURE OF EXPLAINING (NEW) CONCEPTS

Discourse element	Function	Example / Communication strategy used
Introduction of new concept / topic	Speaker introduces a new concept or a new topic (usually by using a technical term or a code-switch)	<i>Do you know XY; Have you heard of XY?</i>
Check level of knowledge	Speakers need to find out whether they have the same level of knowledge or if one speaker is lacking knowledge. Speaker who doesn't know the concept could ask for clarification.	<i>Yes, I've heard of it. No, I don't know what it is. What is XY? I don't know if you know that. I don't know if you've tried XY</i>
Evaluation*	Speaker expresses his/her reaction to (not) knowing the concept	<i>Oh you should try it. / It's so good. / It's really cool / Oh my God, how do I explain it?</i>
Explanation / Elaboration	Speakers share experience, relate to known concepts and compare differences and similarities. Speaker who knows a concept (or introduces TV series) explains the concept to the interlocutor. Speakers verify concepts by asking for clarification or making assumptions.	Spelling (B-R-E-Z-E-L), describing how it is produced / made (<i>it's made with vinegar</i>) , breaking down the process/concept in many steps / pieces of information; linking to prior knowledge (<i>it tastes like cabbage</i>); comparing to known concepts (<i>it looks similar to XY</i>); using non-verbal elements (gestures); using multimodal features (e.g. sending a picture in the chat); describing the story; mentioning main characters / the city it is set in; actors / actresses; genre; reasons for watching it
Result	Speakers come to an agreement (similarity or difference). Speaker who learned about the concept, expresses (dis)like or signals understanding. Speakers who know concepts encourage interlocutors to try it out (recommendation).	<i>Oh, that's interesting. Oh, I might try it. You should try/watch it. That's similar to XY then. I think it's the same, just XY.</i>

This discourse structure is similar to a general sequence of **topic introduction** that consists of the following three stages: (1) introductory, (2) acknowledging and (3) establishing phase (see Geluykens 1993: 188, Brunner 2015). These phases and the sequence in general is characterized by a “collaborative” effort of the participants to organize discourse (Geluykens 1993: 182). Geluyken’s introductory phase is very similar to the first step mentioned in table 42, in which the speakers **introduce a new concept** or a new topic. Steps 2 and 3, **checking level of knowledge** and **evaluating the concept**, is very similar to what is described as the second step in Geluykens’ topic introduction phases, namely the acknowledging phase. In this phase, the participants come to an agreement that they continue the conversation about the topic that had been introduced. In the case of explaining a new concept, the speaker who introduces the topic usually checks if the interlocutor knows the concept or not, and depending on the knowledge, the concept is either elaborated or explained and the following steps are carried out or the topic is changed again. The **elaboration or explanation phase** in table 42 is very similar to the establishing phase, in which the topic (or the concept) is discussed in greater detail. The last step, that is described in table 42, **the result**, means that the speakers come to an agreement about the concept that had been introduced and discussed. This phase is characterized by evaluative language as the concept is evaluated, which means that the speakers usually express like or dislike or recommend that the interlocutor should try the concept. This step depicts the closing of that sequence of explaining a new concept and it is a new step that was not mentioned in Geluyken’s topic introduction phases.

The main phase of the sequence of explaining a new concept is the third step that I called **explanation or elaboration** (see table 42). Diemer and Brunner (2020), who investigated how identity is created and negotiated in conversations about food in ViMELF, found elaborate explanations on food items as a sign of cultural expertise (Diemer & Brunner 2020: 174). They argue that these explanations are similar to the ones in food blogs or recipes and that these explanations function to establish authority (Diemer & Brunner 2020: 174). These “explanations in ViMELF are indeed mostly lengthy” (Diemer & Brunner 2020: 174), which is the same for TaCoCASE as this sequence usually takes 30 seconds or even

longer (see example 19). Besides these elaborative explanations, they find that cultural expertise is characterized by the use of expert vocabulary, recipe telling, food storytelling as well as evaluative statements about food (Diemer & Brunner 2020: 172 f). These aspects can be found in TaCoCASE as well and I will (and have already) draw upon these aspects in the explanations of the several examples from TaCoCASE. I will explain two selected examples that follows the established discourse structure (see table 42) in greater detail in the following.

(19) Discourse structure – Sauerkraut (07SB149BI31; #03:30#)

#	Conversation (intonation units)	Discourse elements
1	SB149: some- some typical,	Introduction of new concept (including check of knowledge)
2	like stereotypical German [uh:m]	
3	BI31: [°yeah°]	
4	SB149: things are good as well. (1.3)	
5	so I don't know if you- if you've ever eaten	
6	Sauerkraut ((German (0.8)))?	
7	or Kraut ((German (0.5)))? ...	
8	BI31: no, what's that?	Check level of knowledge
9	SB149: it's like <cabbage>? {lifts head up} .. A:[nd]	
10	BI31: [O:H.]	Explanation (E1) Comparison
11	SB149: that [is-] {moves hands in circle}	
12	BI31: [°>okay.<°]	(E2) Description of production process
13	SB149: made (up) with some vinegar and stuff like	
14	that,	
15	and it's uh: left inside a big box o:r a- a- a	
16	round (thingy)?	
17	BI31: yeah, [a (thingy)]. {touches shoulder & pulls sleeve}	
18	SB149: [y:nd] where ((laughing))-	
19	where it is- where it stays,	
20	and uh:m slowly {moves hands in circle}	
21	.. turns into Sauerkraut ((German(0.9))),	
22	so: the cabbage is.. slowly turning into that.	
23	.h a[nd u:h.]	
24	BI31: [>Sauer<]. ((German (0.3)))	(E3) Further associations
25	{leans forward & looks to side}	
26	SB149: >yeah<.	
27	.. h ..it's-	
28	BI31: ((chuckles))	
29	SB149: ((hehe)) it's like the- the typical stereo..	
30	{looks to side & raises eyebrows}	
31	typical thing that Americans say about Germans,	
32	is that they are eating Kraut ((German (0.4))).	
33	(1.5) so they also refer to us as the Krauts	

34		((German (0.5))). ...		
35		BI31: o:h it's {leans forward} [OKAY THAT'S		Result
36		[INTERESTING.]		

The participants talk about eating habits and (stereo)typical food. The German speaker **introduces the concept** *Sauerkraut* and at the same time checks whether his interlocutor knows the concept (lines 4-8). SB149 uses a code-switch, in fact the *mots juste*, *Sauerkraut*, at first before providing a translation for her interlocutor, to signal expertise on the one hand but at the same time signaling understanding if the concept is not known. BI31 explains that she does not know the concept (line 8), which is a prompt for SB149 to explain the concept, which he does in the following (lines 9-35). SB149 uses several strategies to **explain** what *Sauerkraut* is and he breaks down the information of the concept in a series of pieces of information: He starts off with a **comparison (E1)** by linking to previous knowledge as he assumes that BI31 knows what cabbage is. He **describes how it is produced or processed (E2)**, supported by gestures (moves hand in circles in lines 11 and 20), which is similar to recipe telling and underpins his cultural expert status (see Diemer & Brunner 2020). He describes that *Sauerkraut* is *made up with vinegar* (13), *left inside a big box* (15), and repeating the actual word (code-switch) *Sauerkraut* (21), which the *cabbage is slowly turning into* (22). It is important to note that in his explanation, he uses vague language (thingy), hedges (*and stuff like that*) and non-verbal elements such as gestures in order to make up for his lack of vocabulary. Even if he is not the linguistic expert, he is assuming cultural expertise by adding some **further information**, actually an association (E3) or an anecdotal evidence, by explaining that Americans sometimes refer to Germans as the *Krauts* (lines 31-33), drawing on *Sauerkraut* being a traditional food in Germany. As a **result**, the interlocutor signals understanding and states that she finds this piece (or the pieces) of information very interesting (line 36). This examples shows very well how the discourse is structured when a new concept is being introduced while only one speaker knows the concept and the concept is unknown to the interlocutor. The steps outlined in table 42 are very nicely drawn upon in this example and it seems that the structure also helps to create clarify meaning and contribute to mutual understanding.

The discourse structure of explaining a new concept also applies to other cultural concepts besides food, such as TV series. Example 20 is an excerpt from a conversation in which the participants talk about a TV series called *Bones*.

(20) Discourse structure – Bones (08BO15BI36; #02:00#)

#	Conversation (intonation units)	Discourse elements
1	BO15: uh:m, #00:02:00-3#	
2	... we own,	
3	and it's like our family .. drama television show,	
4	BI36: yeah.	
5	BO15: uhm it's called <i>Bones</i> ,	Introduction of new concept
6	<i>I don't know if you've heard of it?</i>	Check level of knowledge I
7	BI36: <yes>.	
8	<i>I think,</i>	
9	y- yeah,	
10	I think I,	
11	<i>if I'm thinking of the right one,</i>	
12	<i>then maybe. ((chuckles)) ((snuffles))</i>	
13	BO15: <i>so she's uh,</i>	Explanation I (series)
14	<i>the main character,</i>	(E1) main character(s)
15	<i>there are two main characters,</i>	
16	<i>but the story really revolves around her,</i>	(E2) story
17	<i>and she's a uh anthropologist,</i>	
18	<i>a forensic anthropologist.</i>	
19	BI36: yeah.	
20	BO15: and she <i>helps an FBI agent solve crimes.</i>	
21	BI36: <i>yeah [I think I],</i>	Check level of knowledge II
22	BO15: <i>[in Washington]. #00:02:30-9#</i>	(E3) city
23	BI36: yeah.	
24	BO15: ((ehh))	
25	BI36: <i>I think I've [heard of this].</i>	Check level of knowledge II
26	BO15: (((ehh)) so it's),	
27	BI36: [okay].	
28	BO15: [yes] I,	
29	<i>we really like it,</i>	Evaluation I
30	uh:m,	
31	... so I live with,	
32	it's just me, my mom and my sister,	
33	BI36: [okay].	
34	BO15: [and] <i>there are very strong female leads in the movie,</i>	Explanation II (reason for watching it)
35		
36	<i>so we LOVE [watching it].</i>	Evaluation II
37	BI36: (((ehh)))	
38	BO15: cause they are very strong characters,	
39	and they're kind of [crazy ((chuckling))].	
40	BI36: <i>[that's very] cool though. ((chuckles)) (((snuffles)))</i>	Result
41	BO15: [yeah].	

Example 20 shows how the participants create common ground by **talking about a TV series** that both (supposedly) know. BO15 **introduces the TV series** by stating that she and her family *own* it (note: they probably bought it) and it is the drama show they are watching together as a family: *Bones* (2-5). In fact, this is a very personal note and an individual context as it seems to be a family ritual for BO15 to watch the series together with her family, which is similar to what Brunner and Diemer (2020: 182) described as cultural expertise through individual food context. BO15 checks, whether her interlocutor, BI36, knows the series (6) and BI 36 is a little hesitant in her answer using many hedges (*I think, If I'm thinking of the right one, maybe*), but she supposedly knows the TV series as well. This hesitant answer is a prompt for BO15 to explain what the series is about (13-24). She starts off by talking about the main character(s), *a forensic anthropologist who helps an FBI agent solve crimes* (13-21) and where the series is set in, namely in Washington (22) to underpin her expert status. With these pieces of information, BI36 recalls the series and she states that she thinks she has heard of it (21), which she repeats due to overlap and not wanting to interrupt her interlocutor (25). BO15 **evaluates** the series by stating that she and her family *really like it* (29), which is again a personal preference or evaluative statement. She even adds another **explanation**, but this time not about the series itself, but about her reason for watching it – again signaling expert status and providing personal information. She states that she and her family, which consist of three women, like the strong female characters (34). She repeats her evaluation, this time with an even stronger emphasis, using a stronger verb (*love* instead of *like*) and she put a stress on *LOVE* (indicated by capital letters in the transcript and by loudness in the audio file) while uttering it. As a result, BI36 understands BO15's reasons for watching it and why BO15 really likes *Bones*, by saying *that's very cool though*. This explanation of the concept or TV series follows the general discourse structure that was established in table 42, but it is interesting that this explanation is very much influenced by personal preferences and evaluations. This is something that is sometimes referred to in food discourse (see Diemer & Brunner 2020), but that is not established very much in previous research and should be investigated more closely in the future.

In general, there are more examples that underly the same discourse structure when explaining (new) concepts. This discourse structure is very universal and it is very similar to the phases of topic introduction sequences as investigated by Geluykens (1993) and Brunner (2015). Since the term concept applies to all cultural aspects mentioned in the analysis, including food, customs and traditions as well as TV series, examples from these contexts follow the same structure in TaCoCASE. In fact, the discourse structure could as well be applied to other cultural concepts in intercultural encounters. This discourse structure is very useful in EFL teaching as learners could use this structure and apply it in different contexts when they are in intercultural situations and want to explain cultural concepts (see also case study in section 5.5.2.)

All in all, the **analysis chapter** delivered many findings about linguistic and (inter)cultural aspects in TaCoCASE. It unveiled insights about words and phrases and the language used in TaCoCASE as well as how people exchange cultural concepts that could be useful for the application of the corpus in particular in EFL teaching.

In the **linguistic analysis**, lexical as well as pragmatic elements were discussed in order to provide evidence for the application of the corpus when it comes to establishing communicative competence. Linguistic and pragmatic competence are important aspects of establishing communicative competence (CEFR 2020). In terms of lexis, the most frequent words and lexical bundles in TaCoCASE and in ViMELF had been identified, which are very characteristic of spoken language and that showed that TaCoCASE is very representative in terms of spoken language. These words and phrases (see table 26-27) could be used for teaching as they appeared frequently in the two corpora and are therefore relevant. Patterns for frequent lexical bundles such as *I don't know* were detected and collocates of *really* were investigated before their use in context was explored. Following the lexical approach, the identified lexical bundles and collocates are important to learn as learners need ready-made phrases and fixed expressions to be able to communicate (Lewis 2002; Thornbury 2002; Haß 2006; Hutz 2018).

The examples that were discussed could very well be integrated in EFL teaching and could help learners build on their vocabulary and thus on communicative competence.

In the **lexical analysis**, I also examined differences in language use among NS and NNS. I created vocabulary profiles for the NS and NNS and compared the lexical density and TTR of what the NS and NNS said in order to find differences in terms of lexis. The results indicate that there are no big differences between NS and NNS in TaCoCASE as far as the use of lexical items is concerned. This suggests that the setting, the intercultural ELF setting, has an impact on the language in TaCoCASE and that the speakers should be described as intercultural speakers rather than NS or NNS. This fact provides proof that the intercultural speaker as model in EFL teaching is more suitable than the NS model as the differences in language use were not very large between NS and NNS.

As far as the **discourse marker analysis** is concerned, I found different functions of the frequently used discourse marker *so*: result and inference, action prompt, topic developer and sequencer, connection marker, marker of emergence from incipency and hesitation and floor-holding device, which were also identified in previous research. When investigating collocates of *so*, I found that *okay*, another discourse marker, frequently co-occurs with *so*. This led to the analysis of DM combinations, and in particular the combination *okay so*. When looking at DM sequences I found that the sequence *okay so* is preferred over *so okay* and that the co-occurrence could also appear across speaker turns. Additionally, I found that DM combinations (partially) fulfill different functions than the functions that arise from its components. The majority of the functions (of the DM combination) were combined functions, even though new functions – such as *abbreviating comprehension*, *rapport through summary*, *strengthened intersubjectivity* and *strengthened topic start* – and functions that arise from one of the components also appeared. Again, examples were taken from TaCoCASE and discussed in order to illustrate the different functions. These findings provide useful information in terms of establishing pragmatic competence in the EFL classroom when it comes to teaching discourse markers. Acquiring discourse markers, looking at their form and meaning as well as their function and how they are used contributes to establishing pragmatic competence, which is an important aspect of

communicative competence (CEFR 2020). The examples could be used directly in the EFL classroom in order to raise awareness for form, function and meaning of discourse markers.

Finally, I looked at differences in terms of discourse marker use among NS and NNS. The results indicate that there is no difference in terms of functions of *so* between NS and NNS and that the results are quite in line with previous research and a previous study we conducted with ViMELF (Collet et al. 2021). As far as DM combinations are concerned, no difference for sequence preference among native and non-native speakers could be stated. The results indicate that there is no big difference in DM use between NS and NNS, which underpins the hypothesis that the language of NS and NNS is not very different in TaCoCASE. This is probably due to the intercultural ELF setting and provides proof for the concept of the intercultural speakers in TaCoCASE. It also supports the hypothesis that the intercultural speaker is a more suitable model in EF teaching than focusing on NS only (see discussion in section 2.1.5.2.7.).

As far as the **(inter)cultural analysis** is concerned, the results provide important aspects that could be used in the EFL classroom when establishing intercultural competence. From analyzing TaCoCASE, it can be stated that cultural aspects play an important role in the conversations in TaCoCASE as cultural aspects were found in 2/3 of the conversations – more than originally expected from the topic prompts. A semantic analysis of food-related words derived a semantic profile of words that are used in TaCoCASE in context with food, such as food items, verbs that are used in context with food as well as locations that have to do with food and others (see Rüdiger & Mühleisen 2020; Diemer & Diemer 2020; Santino 1996; Rutagand 2024). Taking a closer look at the content of the conversations, it was examined what the participants talked about in terms of food, customs and traditions and popular culture. I identified elements that the participants discussed, such as stereotypes, traditional food, festivities and holidays, sports and TV series, and I explained and discussed these elements with regard to mediation strategies and negotiation of meaning. These findings could be used in the EFL classroom when establishing intercultural competence. Having examples for cultural topics and words related to cultural topics is important when talking about cultural aspects. If learners know for example what

they could talk about in terms of food – food items or locations where food is eaten and have food-related words at stake – it will be easier for them to talk about cultural issues, the content is more authentic and through talking about cultural aspects and mediating concepts, they establish intercultural competence.

In a further step, I investigated how the participants talk about cultural aspects. Mediation and exchanging (new) cultural concepts is very important when it comes to intercultural communication. I was interested in the strategies for mediating (new) cultural concept the speakers used. If the concept is known to both speakers, I found that the participants in TaCoCASE shared their experience with the concept or compared the respective concepts they know. If the concept was only known to one speaker and the concept was from the speaker's own culture, the speaker explained the concept to clarify meaning, which is usually related to a cultural expert status. If the concept was from the interlocutor's culture, the speaker asked for clarification or tried to verify his or her concept, assuming that the interlocutor is the cultural expert and can explain the concept. It is important to note that this distinction own culture versus interlocutor's culture only counts for elements that are culture-specific and national concepts. Transcultural phenomena or elements would not call for this distinction and the participants could still explain a concept or ask for clarification. Especially TV series and other aspects of popular culture fall into this category and one participant very nicely stated the following, after a long conversation about TV series and movies the participants watched and liked: *"I think it's so cool, how we like the same things, but they're not our own culture essentially. That's pretty cool."* (BI36; 08BO15BI36, 45:15). The findings from the discourse analysis provide many strategies that learners could use when talking about cultural concepts themselves. The analysis also suggests that TV series have great potential to be used and analyzed in the EFL classroom in terms of talking about popular cultural aspects. In general, the examples that I examined make clear that both – culture-specific as well as transcultural phenomena – should be discussed in the EFL classroom.

When mediating (new) concepts, I identified a basic pattern in terms of discourse structure that consists of the following elements: introduction of (new) concept, check of knowledge, evaluation,

explanation / elaboration and result, which is very similar to introducing a new topic (Geluykens 1993; Brunner 2015). Sample expressions for each of the steps were provided so that the strategies can be (re)used in different contexts. Two examples from the conversations were examined very closely and analyzed with regard to the identified discourse structure. The cultural analysis shows very clearly how common ground is created and how the participants exchange cultural concepts, which contributes to mutual understanding and establishing intercultural competence. The findings indicate that there is great potential in looking at conversations about cultural topics in the EFL classroom. It is not only the content – what aspects could be talked about – that the conversations offer. It is also the way people negotiate meaning and create common ground and how they evaluate cultural aspects. This offers new perspectives in establishing (inter)cultural competence as discursive elements and attitudes are included, which are usually not very much discussed in intercultural learning (see König et al. 2022). Using examples from TaCoCASE that were presented in the analysis to establish intercultural competence thus suggests a new approach in intercultural learning – discursive intercultural learning (Hallet 2008; König et al. 2022).

It becomes clear that the analysis and the results presented in this chapter provide a **firm basis** and many examples of what is **relevant and should be taught at school** in terms of linguistic as well as (inter)cultural aspects. The examples from TaCoCASE are a rich source for teaching and the application of TaCoCASE in the classroom. Since the idea is to teach only aspects that are frequent and that are empirically proven, this chapter provided valuable and relevant results. The chapter investigated language use as it is, which means that authentic words and phrases as well as common expressions were identified. These authentic results should be used in the classroom so that language teaching is more authentic and in return, the language use by the learners becomes more authentic as well. The results from the intercultural analysis provide **relevant (cultural) topics** that could be discussed in the EFL classroom. Additionally, the (inter)cultural analysis illustrated how the speakers explained new concepts and provided strategies the speakers used and I identified a **common discourse structure** for explaining (new) cultural concepts that could be very useful for the learners.

Thus, examples from TaCoCASE are very well suited to teach (inter)cultural aspects and raising cultural awareness among learners in discursive intercultural learning, which is demanded and suggested as new way of establishing ICC in the EFL classroom (see Hallet 2008; König et al. 2022). The methods that I applied provide a **new approach** of finding aspects that are relevant for teaching. The linguistic as well as (inter)cultural findings also suggest new topics in EFL learning, for example new words and phrases that should be taught, integrating TV series in intercultural learning, talking about cultural issues in discursive intercultural learning. Since the analysis also provided proof that the intercultural speaker exists in TaCoCASE and the language use between NS and NNS is not very different, it can be stated that the analysis contributes to the discussion of the model change in EFL teaching towards the intercultural speaker. In general, it becomes clear that through applying corpus methods and finding new and relevant linguistic as well as (inter)cultural aspects, a mindset change in EFL teaching is proposed. The following chapter will focus on the application of these findings, and the corpus in general, in the EFL classroom.

5. APPLICATION: TACOCASE AND EFL TEACHING

After having analyzed linguistic and cultural aspects in TaCoCASE in the previous chapter, the focus in this chapter lies on the **application of these evidence-based findings in the EFL classroom**, and in general, on how corpora – in this case TaCoCASE – can be used in the EFL classroom. While chapter 4 looked at the analysis from a (corpus) linguistic and (inter)cultural point of view, this chapter looks at the applications of the corpus from an applied linguistics and EFL teaching perspective. The **aim of this chapter** is twofold: (1) It will illustrate how examples from TaCoCASE meet the CEFR guidelines to establish communicative and intercultural competence and (2) it will explain how TaCoCASE can be used in the EFL classroom to support a modern way of EFL teaching. This chapter brings together the main findings from the previous chapter and applies them in an EFL context. This is the main contribution of this dissertation as new material is created and new methods are applied, which promotes a mindset change in EFL teaching.

It would go far beyond the scope of this work to look at every aspect mentioned in the CEFR and provide examples from TaCoCASE for the different competences mentioned in the CEFR. Therefore, a selection based on the findings of the analysis in chapter 4 and the special features of TaCoCASE was chosen that will be outlined in the following. The first part of this chapter takes the **CEFR as a basis** as it dives into two main aspects of language proficiency as described by the CEFR – **communicative language competence** and **communicative language strategies**. As for communicative language competence, the focus will be on linguistic and pragmatic competence. Communicative strategies such as interaction and mediation strategies help learners to communicate successfully with interlocutors in order to be successful users of a language. Excerpts from TaCoCASE can serve as good examples for teaching interaction and mediation strategies. As far as intercultural competence is concerned, some aspects will be covered among mediation strategies already. Additionally, the operators for building pluricultural repertoire will be taken as a basis and examples from TaCoCASE will be assigned to them in order to find ways to establish intercultural competence.

Due to its multimodal set-up, TaCoCASE provides examples for aspects of **online interaction** and **non-verbal communication**. These multimodal features in TaCoCASE play an important role in interaction and they are usually not highlighted very much in the EFL classroom (Byram 2021; see also section 5.4.2.). Online interaction is a fairly new topic in teaching – even though it is very present in our everyday life – and has only been mentioned in the new version of the CEFR (2020). The amount of teaching material already existing is very limited. Non-verbal communication is a key feature of interaction and multimodal meaning-making in a conversation and has not gained much attendance in EFL teaching, yet (Byram 2021: 36).

Going even further beyond frameworks and the curriculum, I will demonstrate how selected examples from TaCoCASE can be used in the EFL classroom by providing some possible teaching tasks. In a **case-study**, two lessons will be described in detail that cover the following topics: (1) teaching vocabulary with TaCoCASE to establish communicative competence and (2) mediating (new) concepts to promote intercultural competence. Finally, the last section highlights the benefits of TaCoCASE in an EFL teaching context.

As mentioned in chapter 3.2., I established **learning objectives** for each competence in order to describe what the learners should be able to do at the end of each chapter (sections 5.1. – 5.4.). The learning objectives are based on the competences in the CEFR and are formulated in a can-do approach, which emphasizes the learner-centeredness (see section 2.1.4.). They follow Bloom's taxonomy of learning objectives and thus draw on different cognitive levels (see also section 2.1.1.). They mainly serve two functions: (1) They structure each chapter by setting an aim for each chapter and (2) they draw on the competences that are mentioned in the CEFR, which marks the syllabus fit of the corpus and shows to what extent the extracted examples and the teaching tasks establish the demanded competences. Since learning objectives are also used in teaching and in curricula, I wanted to make this chapter as authentic and as applicable as possible. The objectives could thus be directly integrated in syllabi or lesson plans, which is one of the main benefits of this approach and the application chapter in general.

I will provide several **teaching tasks** throughout this chapter that could directly be used in the EFL classroom. Authentic examples and concrete teaching tasks are often demanded by applied corpus linguistics, but still not many exist (see Römer 2011; Mukherjee 2002; O’Keeffe et al. 2007; see also section 2.2.2.). I try to close this gap by presenting the teaching tasks for various applications, including communicative aspects as well as intercultural and multimodal aspects. While communicative aspects are usually drawn upon in examples using corpora, the latter are very rare, which makes the teaching tasks and the contribution in this work even more valuable. The teaching tasks provide authentic tasks and are developed according to Bloom’s educational objective taxonomy (see figure 2), which means that they are designed to address different levels of Bloom’s taxonomy and therefore draw on different cognitive domains and levels of requirements (see section 2.1.1.). In general, this chapter is designed to be very authentic and applicable in the EFL classroom, using learning objectives and providing concrete teaching tasks, that could be used in the EFL classroom. It shows how examples from TaCoCASE match the CEFR requirements and thus how well TaCoCASE and corpora in general can be used to establish the required competences, in particular communicative and intercultural competence.

5.1. TACOCASE AND COMMUNICATIVE COMPETENCE

Communicative competence has already been discussed at length in section 2.1.5.1. and the CEFR describes communicative competence as follows: “Communicative language competences are those which empower a person to act using specifically linguistic means” (CEFR 2001: 9). (Functional) communicative competence is very dominant in language teaching in Germany (Decke-Cornill et al. 2015: 161) and is a major competence to establish. According to the CEFR, communicative competence embraces **linguistic competence**, **pragmatic competence** and **socio-linguistic competence** (see section 2.1.5.1.1.). The focus in this section lies on establishing communicative competence, in particular linguistic and pragmatic competence. It will be illustrated how linguistic competence can be

established through teaching vocabulary such as single words, collocations and n-grams. As far as pragmatic competence is concerned, examples with regard to discourse competence and the use of discourse markers will be analyzed and explained. These applications are based on the findings in the previous chapter (see chapter 4). The examples and methods can be used in the EFL classroom in a learner-centered inductive way of teaching in order to establish linguistic and pragmatic competence and thus communicative competence. The examples depict concrete situations for the objectives mentioned in the education frameworks, the CEFR in particular, and provide evidence of how actual speakers used linguistic and pragmatic language features in actual conversations. The examples depict authentic language in use and are thus very authentic and facilitate a communicative teaching approach.

The objectives that should be reached by the learners in this chapter in terms of communicative competence are the following:

Objectives for communicative competence:

1. The learners acquire new words and phrases. They can describe linguistic aspects of words (lexis, grammar, phonological aspects) and can use these words in context. They can name words and phrases that organize and structure discourse (discourse markers) and they can describe how they are used with their respective functions.
2. The learners can use TaCoCASE and corpus methods in order to find out about linguistic and pragmatic aspects of words.
3. The learners can evaluate their language use (linguistic and discourse competence) with regard to actual language use in TaCoCASE.

The following sections will focus on how to improve linguistic and pragmatic competence by using examples from TaCoCASE.

5.1.1. IMPROVING LINGUISTIC COMPETENCE

Linguistic competence, as defined by the CEFR, has been already described in greater detail in section (2.1.5.1.1.). There are many aspects which should be taken account when establishing linguistic competence, such as lexical aspects, grammatical, phonological or orthographic aspects. Establishing linguistic competence is thus closely related to **teaching vocabulary**, which I will briefly explain in the following paragraph.

Teaching vocabulary

Acquiring new words and teaching vocabulary is in EFL teaching closely connected to the concept of the **mental lexicon**. The mental lexicon is a complex cognitive archive in our brain, which is assumed to be structured and organized in different categories or nets (Hutz 2018: 138 and Thornbury 2002: 16; see also Decke-Cornill et al. 2015; Haß 2006). It can be described as “the human word store, i.e. that part of our long-term memory where the words are linked systematically in terms of conceptual, semantic and phonological similarity” (Hutz: 2018: 138). Thus, we “store and retrieve words” from our mental lexicon (Hutz 2018: 138) when we are speaking or listening or taking part in an interaction or mediation activity. The way words are stored, namely “highly organized and interconnected” (Thornbury 2002: 16) has an influence on how fast learners can retrieve and produce words. The more structured and organized the presentation of a new word in the learning process is, it is easier for learners to acquire and “know” a word (see Hutz 2018; Decke-Cornill et al. 2015; Haß 2006; Thornbury 2002).

What does it mean to “**know**” a word? Vivian Cook, an applied linguist and second language researcher, stated that “[a]cquiring a word is not just linking form with translated meaning [...]. It is acquiring a complex range of information about its spoken and written form, its way it is used in grammatical structures and word combinations, and diverse aspects of meaning” (Cook 2008: 51). Knowing a word has many dimensions and is not restricted to translating the meaning of a word. Frank Haß (2006: 116; my translation), who is a very renowned German English-subject didactician and

author of the very wide-spread introductory work called “Fachdidaktik Englisch” that is usually used in EFL teaching seminars at universities, provides a whole list of questions and thus **criteria for knowing a word** such as:

- Do I recognize the word in spoken or written form?
- Can I spontaneously use the word?
- Do I know the whole scope of meaning of a word?
- Do I know all grammatical forms of a word?
- Can I pronounce the word?
- Can I write the word correctly? (orthography)
- Am I aware of syntactic restrictions (position in a sentence, collocations)
- Can I differentiate the word from other semantically related words?
- Do I know which register the word belongs to?
- Do I know connotative and emotional meanings of a word?

This is a very helpful list for learners in order to establish the whole (cognitive) depth of a word (Thornbury 2002: 25). It means learners should know several aspects besides meaning of the word, such as syntactic, phonological or orthographic information. It is also important for teachers to know about these criteria in order to select words that are worth teaching. Matthias Hutz (2018: 144), a well-known German English-subject methodologist and expert on second language acquisition and teaching vocabulary and grammar, names frequency as one important criterion on selecting words that are worth teaching and he emphasizes the benefits that corpora can provide. He states that “[i]t is obvious that it is useful to select those lexical items which are frequently used” (Hutz 2018: 144). He even names corpora as good source for frequency information of a word by stating that “[l]inguistic corpora can provide very good surveys of words which are used frequently” (Hutz 2018: 144).

Besides a good **selection of words**, the way teachers present vocabulary also has an impact on the learning process. Due to the net-like structure of the mental lexicon (Hutz 2018; Thornbury 2002; Decke-Cornill et al. 2015; Haß 2006) and due to individual differences and preferences by the learners, it is necessary to present words in many different ways so that the learners make many connections between the words in the mental lexicon (see Thornbury 2002: 30; see also Haß 2006; Decke-Cornill et al. 2015).

According to Haß (2006: 119) vocabulary can be presented through

- the use of real objects
- pictures
- imitation
- using the word in context
- definitions
- derivatives
- paraphrasing
- synonyms or antonyms
- translation

Coming from an applied corpus-linguistic perspective, O’Keeffe and colleagues (2007) suggest **lexico-grammatical profiles** for teaching vocabulary and thus suggest using corpora when teaching vocabulary. They combine corpus linguistics and applied linguistics and suggest these profiles that describe “typical contexts” of the words based on concordance lines and follow a corpus analysis strategy (O’Keeffe et al. 2007: 14). The profiles combine form and meaning of a word and thus provide word lists and the context of use at the same time. O’Keeffe and colleagues suggest criteria for lexico-grammatical profiles, which are similar to Haß’ criteria of knowing a word, but coming from an applied corpus linguistics’ perspective:

Lexico-grammatical profiles (O’Keeffe et al. 2007: 14):

- Collocates
- Chunks and idioms
- Syntactic restrictions
- Semantic relations
- Prosody

Looking at these criteria makes clear that combining form and meaning and using words in context is very important. It is therefore not advisable to only teach single words but teach chunks that consist of the “new” word (see Thornbury 2002; Lewis 2002). This raises the question why it is worth teaching lexical chunks and which chunks should be selected and presented to the learners.

The lexical approach and the importance of teaching chunks

Michael Lewis, a very influential applied linguist and international lecturer of English studies and methodology, who is particularly known for coining the widely used “lexical approach” (Lewis 2002), was the first to challenge the dichotomy between vocabulary and grammar (see Hutz 2018: 142). Ute Römer (2011: 216), a very renowned applied corpus linguist, also suggests a “shift from [teaching] individual words or empty grammatical structures to phraseological items (formulas, n-grams, lexical bundles)” (see also Biber et al. 2004; Hoey 2005; Meunier & Granger 2008; Römer 2011; Sinclair 1991; Stubbs 2001). Thus, words should be taught in context and in larger chunks and “information about patterns [should be] incorporated in language teaching materials” (Hunston & Francis 2000: 272). This principle is crucial to Michael Lewis’ lexical approach, in which he states that “language is grammaticalised lexis, not lexicalised grammar” (Lewis 2002: iv). Following the lexical approach means that lexis is important in meaning making and pre-fabricated chunks are easy for students to use and to include in their language repertoire. Pre-fabricated chunks that learners acquire can again be used as “text building blocks” (Biber et al. 2002: 442), which contributes to coherent discourse (Hyland 2008: 6), increases fluency and leads to a higher communicative output by the learners in general. Chunks and in particular collocations and lexical bundles have been defined in section 4.1.3. and they have been subject to research particularly in corpus linguistics. The main two reasons are that (1) lexical chunks are easy to retrieve with the help of computer programs and corpus tools and that (2) these corpus tools provide statistical evidence and probability information of words occurring together and not just relying on native speakers’ intuitions.

When **teaching vocabulary**, it is usually suggested to start teaching single words and continue with words and phrases (see Decke-Cornill et al. 2015; Thornbury 2002). Usually, for beginner learners, acquiring vocabulary consists mostly of acquiring single words and easy fixed expressions such as greetings or other every day expressions like “Merry Christmas” or “Good luck”, common prepositional phrases like “on the first of October” or compounds like toothbrush or car park (Ellis 1997). The more advanced the learners are, the more complex the chunks become. Since lexical bundles or collocations

can be semantically empty, a higher degree of competence is needed in order to fill these combinations.

Ellis (2008) and Holme (2009) use a **construction schema** (see Liu 2012: 28):

- Filled (idioms), e.g. jump the gun
- Partially-filled, e.g. make somebody to do something
- Unfilled (abstract schema): V + NP + NP

Ellis (1997: 13) argues that language learning consist of **item learning**, for example whole expressions like greetings, and system learning, where learners figure out how expressions are constituted. Thus, unfilled constructions are more on the advanced level whereas filled constructions such as idioms can usually be taught at an intermediate or even beginner level. Ellis' categorization is just another and yet very similar categorization as Biber's (2009) definition of **multi-word sequences** (see section 4.1.3.). The concept and the implications for teaching remain the same.

Following the principles of the lexical approach, it is clear that words should be taught in context and that lexical bundles or **chunks are extremely useful for learners**. First of all, they should not be considered "an optional extra, since the meanings they create are extremely frequent and necessary in discourse, and are fundamental to successful human interaction" (O'Keefe et al. 2007: 75). They also fulfill pragmatic functions and play an important role in organization and management of discourse. They are crucial on an interpersonal level for the speaker-listener relationship and for interactive meaning-making in everyday conversation (O'Keefe et al. 2007: 75). Pre-fabricated chunks can be used effortlessly and leave capacity for other cognitive processes while processing language (Girard & Sionis 2004), which means that learners can produce language faster if they can rely on ready-made chunks instead of having to think about how to construct the correct chunks. These aspects combined with the fact that chunks are usually single intonation units, leads to more fluency in speech. If the chunks are memorized in a correct way and can be retrieved easily, they also contribute to accuracy as the probability for using a wrong preposition with a special item decreases.

In sum, teaching vocabulary and communicative grammar, based on the lexical approach, is state of the art in communicative language teaching (see Decke-Cornill et al. 2015; see also Thornbury 2002; Lewis 2002). With the help of corpus examples, inductive bottom-up language acquisition is possible and the learners acquire new words, expand their vocabulary range and feed their mental lexicon. Teaching and acquiring chunks plays an important role in establishing speaking, listening and interaction skills as they support learners in producing language more correct, proficient and faster. Chunks thus help building communicative competence and are particularly worth teaching.

DDL and establishing digital competence

Teaching vocabulary and establishing lexico-grammatical profiles, is very well suited for using a DDL approach, which is a very learner-centered method and a communicative teaching approach (see section 2.2.2.3.). Mukherjee (2002), a well-known German applied corpus linguist, for example suggested a **learner-centered corpus browsing activity**, that consist of the following steps:

Learner-centered corpus browsing (Mukherjee 2002: 168; my translation):

- Interest in a particular English form (e.g. word, phrase in a given text)
- Using corpus data in order to get more information about its usage
- Drawing conclusions from the corpus data
- Documenting the results of the corpus analysis and presenting them in class

The learners use corpora in order to conduct research on words on their own, which can be classified as DDL (see section 2.2.2.3.). They use corpora to find out about a meaning of a word, a word form or anything that is of particular interest to them. Thus, the aim is to find out the meaning and use of a word, which draws on establishing communicative and linguistic competence. The method, however, is corpus-driven (Tognini Bonelli 2001: 177) as the learners make use of the corpus and corpus tools in order to state and proof hypothesis. In conducting the corpus search the learners establish digital competences as well (see section 2.1.6.3.). The whole DDL activity including the corpus search draws on digital competences as described by the KMK like “searching, processing and storing,

communicating and cooperating, producing and presenting, solving problems (and taking action) and analyzing and reflecting” (KMK 2016: 15). DDL activities thus not only promote communicative competence but also digital competence.

Linguistic competence in the CEFR

Teaching vocabulary belongs to linguistic competence and communicative competence in general. Communicative and linguistic competence have been defined in section 2.1.5., already. The key aspects of what the CEFR defines as linguistic competence are described in the following:

Scales of linguistic competence (CEFR 2020: 131)

- General linguistic range
- Vocabulary range
- Grammatical accuracy
- Vocabulary control
- Phonological control
- Orthographic control

These scales are very similar to Haß criteria of knowing a word or what Thornbury describes as (cognitive) depth of vocabulary knowledge mentioned and discussed earlier in this section (Haß 2006; Thornbury 2002). Teaching vocabulary, choosing words and presenting them to the learners draws on three aspects mentioned by the CEFR: (1) general linguistic range, (2) vocabulary range and (3) vocabulary control. These aspects are further described in the CEFR by key concepts, which are operationalized so that they can be assessed (see section 2.1.1. and section 2.1.5.1.1.).

The following figure presents the three selected key concepts that the CEFR described to establish linguistic competence and that are of interest when teaching vocabulary.

Key concepts operationalized for linguistic competence (CEFR 2020: 131):

General linguistic range	Vocabulary range	Vocabulary control
• range of settings – from A1 to B2, then unrestricted; • type of language: from memorised phrases to a very wide range of language to formulate thoughts precisely, give emphasis, differentiate and eliminate ambiguity; • limitations: from frequent breakdown/misunderstanding in non-routine situations to no signs of having to restrict what they want to say.	• range of settings – from A1 to B2, then unrestricted; • type of language: from a basic repertoire of words/signs and phrases to a very broad lexical repertoire including idiomatic expressions and colloquialisms.	• familiarity of topics (A1 to B1); • degree of control (B2 to C2).

FIGURE 52: KEY CONCEPTS OPERATIONALIZED FOR LINGUISTIC COMPETENCE (CEFR 2020: 131)

General linguistic range has to do with the vocabulary at the learners' disposal in general and it also comprises complex language use and the ability to take risks and testing out and learning new words (CEFR 2020: 130). Learners should be able to "express themselves clearly without much sign of having to restrict what they want to say" (CEFR 2020:130) on B2 level. Vocabulary range is concerned with the "breadth and variety of expressions used" (CEFR 2020: 131). On B2 level, for example, learners should understand and use "technical terminology" as well as have a "good range of vocabulary" of general topics, they should be able to "produce appropriate collocations" and understand and use "specialist vocabulary" (CEFR 2020: 131). As far as vocabulary control is concerned, the learners should be able to express themselves appropriately. This means that on B2 level, learners should have a high "lexical accuracy" (CEFR 2020: 132). In order to establish linguistic competence, these scales and key concepts should be addressed and teaching vocabulary has to follow these requirements. The following teaching tasks help establishing linguistic competence, having these scales and key concepts in mind and drawing on the data and the analysis in the context of this work.

The following teaching tasks provide made-up examples on how vocabulary could be taught with the help of TaCoCASE and DDL.

Teaching Tasks

(1) Teaching Tasks: Vocabulary I

Drawing on the implications for teaching discussed in this chapter and the findings in section 4.1. and in particular 4.1.1., it is useful that the learners acquire the 30 most frequently used words in TaCoCASE (see table 20). One example of the 30 most frequent words that has been analyzed is *really* (see section 4.3.). The following tasks present made-up examples how the word *really* could be taught in the classroom. These tasks are learner-centered DDL activities that could be completed by the learners with the help of a corpus tools.

Task 1: Find words (collocates) that are used together with *really*.

Task 2: *Really* is often used with adjectives or adverbs. How is *really* used in combination with these words and in sentences? (Syntactic restriction)

Task 3:

Go to WebCorpLSE and search for the word *really* in TaCoCASE. Find out what *really* means and how it is used. Create a vocabulary profile for *really*.

Task 1 and 2 focus on the word *really* and how the word is used in combination with other words or in a sentence. Task 1 is concerned with collocations, which means words that are frequently used in combination with *really*. Task 2 is a follow-up task and expands the scope from the single word and collocate of that word to the whole sentence. These tasks can be completed with corpus tools and with a collocation analysis and a concordance analysis in AntConc for example (see section 2.2.1.3.). The corpus tools provides statistical evidence of frequently used words in combination with *really* (see section 4.1.3.). The learners could look at the collocate list (see table 26) and compare the words and the word forms of the different collocates. In a second step, they could sort the results by statistical values and frequency, so that they find the most frequent and strongest collocates of *really*. As the results in section 4.1.3. indicate, *really* is used in combination with adjectives and verbs, usually as modifier in front of adjectives and adverbs (for example *really nice*, *really big*, *really cool*). In terms of function, it used to express emphasis or truthfulness. The learners could look at the whole

concordance in a concordance line analysis (task 3) and find out how *really* is used in a sentence and whether there are any syntactic restrictions.

Task 3 is a learner-centered DDL activity that draws on the concept of lexico-grammatical profiles of words (see O’Keeffe et al. 2007: 14). This example needs some practice beforehand, which means that learners must know how to use the corpus search tool, how they read concordances and how they can retrieve information from their research in order to create a lexico-grammatical profile or a vocabulary profile. They should also know what elements and information a lexico-grammatical profile involves, which means that this example can only be used in advanced stages of teaching. A similar example will be discussed as case study in section 5.5.1.

The following task (task 4) is also concerned with collocates of *really*, similar to task 2, but in this task, WebCorp and TaCoCASE are used as tool and database.

Task 4: Look at collocates of *really*. Find collocates with a similar meaning and try to figure out differences in use.

In task 4, the learners should find collocates of *really* and focus on collocates with a similar meaning but different uses. The learners could conduct a collocate search in WebCorp for *really* (as hands-on activity) themselves. Alternatively, the teacher could present the collocate list of *really* (as hands-off activity).

CASE / ViMELF: Video-Mediated English as a Lingua Franca					Sub-corpus: TaCoCASE
					Words 109,541
					Documents 15
					Advanced Search
Search Compare really: Concordances Collocates					
Span 1 Left & Right Words excluding stopw Ordered by Log-likelihood Generate Collocates					
Word	Left Freq.	Right Freq.	Total	Score	
nice	0	33	33	200.213	
really	25	25	50	169.766	
cool	0	15	15	65.595	
big	1	11	12	49.852	
good	0	14	14	39.320	
oh	21	1	22	38.132	
interesting	1	9	10	35.550	
interested	0	5	5	32.286	
bad	0	7	7	28.724	

FIGURE 53: RESULTS FOR COLLOCATES OF *REALLY* IN WEBCORP

Figure 53 displays the results for collocates of *really* sorted by log-likelihood (collocate strength). The learners could for example choose the words *nice*, *cool* and *interesting* as words that are frequent collocates of *really* and that have a similar meaning.

WebCorp allows to **compare two words** with each other, so the learners could try out the different combinations: *nice – cool*; *cool – interesting* and *nice – interesting*.

TABLE 43: SCREENSHOTS FROM WEBCORP - COMPARISONS

<i>nice – cool</i>	CASE / ViMELF: Video-Mediated English as a Lingua Franca Sub-corpus: TaCoCASE Words 109,541 Documents 15 Advanced Search		
	Search Compare really: Concordances Collocates nice ↔ cool Span 4 Words excluding stopw Generate Collocates Collocate with <i>nice</i>, but not <i>cool</i> talking evening Collocate with both <i>nice</i> and <i>cool</i> really quite Collocate with <i>cool</i>, but not <i>nice</i> okay		

<i>cool – interesting</i>	CASE / ViMELF: Video-Mediated English as a Lingua Franca Sub-corpus: TaCoCASE Words109,541 Documents15 Advanced Search Search Compare really: Concordances Collocates cool ↔ interesting Span 4 Words excluding stopw Generate Collocates Collocate with cool, but not interesting Collocate with both cool and interesting Collocate with interesting, but not cool pretty quite actually oh sounds kind really
<i>nice – interesting</i>	CASE / ViMELF: Video-Mediated English as a Lingua Franca Sub-corpus: TaCoCASE Words109,541 Documents15 Advanced Search Search Compare really: Concordances Collocates nice ↔ interesting Span 4 Words excluding stopw Generate Collocates Collocate with nice, but not interesting Collocate with both nice and interesting Collocate with interesting, but not nice talking quite kind nice really evening

The expectation horizon (“Erwartungshorizont”) that I had in mind when designing the task is the following:

- Similarities: The learners find that *quite* and *really* are collocates of *nice*, *cool* and *interesting*, which means that the following combinations of words appeared in the corpus: *quite nice*, *quite cool*, *quite interesting* and *really nice*, *really cool* and *really interesting*.
- Differences: *Nice* can be used in combination with *evening*, but *cool* and *interesting* are commonly not used in combination with *evening*. It is common to say *nice evening*, but not *cool evening* or *interesting evening*. There are different lexical chunks for different collocates: *have a nice evening*, *it was nice talking to you*, *kind of interesting*, *that’s actually interesting*, *that sounds cool/interesting*. *Pretty* is a frequent collocate of *cool*, but not of *interesting*, which means that the speakers in the corpus preferred to say *that’s pretty cool* over *that’s pretty interesting* (note: both chunks exist; the difference lies in frequency of use).

The following teaching tasks focus on lexical bundles, which means that the scope is not on single words or collocates (combination of two words) any more, but on multi-word units (or lexical bundles or chunks). As discussed at the beginning of this section, lexical bundles are very useful in teaching as they can be used as text building blocks by the learners, which reduces errors and increases fluency.

(2) Teaching Tasks – Lexical bundles

Task 1: Complete the following lexical bundles and give two examples each:

I can't think of....

I don't know ...

Task 2: Think of other lexical bundles that are used in everyday language and write down a short conversation using a lexical bundle.

Task 3: Explain what lexical bundles are and to what extent they can be helpful in language learning?

Task 1 draws on the frequent lexical bundles *I can't think of* and *I don't know* that were found in TaCoCASE (see section 4.1.3.). It is gap filling exercise that the learners could complete very fast by looking at the concordances of these n-grams in WebCorpLSE or using *AntConc*. The learners will come up with many examples that the tool provides from authentic interaction in TaCoCASE and they could select two examples as demanded in the task. As a follow-up, the learners could look at the different patterns that are used with these lexical bundles like, for example *I don't know how*, *I don't know why*, *I don't know if*, so that they can use these bundles in their own speaking or writing. This task could be used at a beginner level – provided the learners are used to working with corpus tools. The level of requirement corresponds to level 1-2 in Bloom's learning taxonomy, while task 2 and 3 are more intermediate tasks, which correspond to levels 5 and 6 in Bloom's taxonomy (see section 2.1.1.; table 1).

In sum, the objectives and tasks presented in this section are designed to build communicative competence and linguistic competence in particular, and are based on the findings in the linguistic analysis (see sections 4.1.1. and 4.1.3). The tasks range from a beginner level to an intermediate level and are concerned with teaching vocabulary in a learner-centered way in a DDL and communicative

teaching approach. The tasks start by teaching single words and their meaning before focusing on combination of words like collocations and multi-word-units such as lexical bundles. The tasks fulfill the requirements in the CEFR and correspond to the scales and key concepts operationalized for linguistic competence in particular (see figure 52), which means that they are relevant for teaching and should be used in the classroom.

5.1.2. IMPROVING PRAGMATIC COMPETENCE

Another aspect of communicative competence besides linguistic competence is pragmatic competence. One aspect of pragmatic competence as per the CEFR is **discourse competence**; besides functional competence and design competence (see section 2.1.5.1.1.). Discourse competence, which has already been discussed in section 2.1.5.1.1. in greater detail, is concerned with how texts or utterances are “organized, structured and arranged” (CEFR 2020: 138). Discourse competence can be established or measured according to the following criteria:

Scales for discourse competence (CEFR 2020: 138 f.)

- Flexibility
- Turntaking
- Thematic development
- Coherence & cohesion

The CEFR lists the following key concepts for establishing discourse competence.

Key concepts for discourse competence operationalized (CEFR 2020: 138):

Flexibility	Turntaking	Thematic development	Coherence & Cohesion
• recombining learnt elements creatively (especially lower levels); • adapting language to the situation and to changes of direction in conversation and discussion; • reformulating points in different ways to emphasise points, express degrees of commitment and confidence, and avoid ambiguity.	• initiating, maintaining and ending conversation; • intervening in an existing conversation or discussion, often using a prefabricated expression to do so, or to gain time to think.	• telling a story/relating a narrative (lower levels); • developing a text, expanding and supporting points appropriately, for instance with examples; • developing an argument (especially B2 to C1).	• linking elements, mainly with logical and temporal connectors; • using paragraphs to emphasise text structure; • varying the types of cohesive devices used, with fewer “clunky” connectors (C levels).

FIGURE 54: KEY CONCEPTS OPERATIONALIZED IN THE CEFR FOR DISCOURSE COMPETENCE (CEFR 2020:138)

Flexibility, as described by the CEFR, is “the ability to adapt language learnt in new situations” as well as formulating thoughts (CEFR 2020: 138). Learners should be able to recombine elements that they learned, adapt language to the situation that they are in and they should be able to formulate points and thoughts as well as to reformulate them in order to make emphasizes or evaluate their utterance (CEFR 2020: 138). Turntaking means primarily taking the “discourse initiative” by opening a conversation, holding the floor and closing a conversation as well as interrupting and actively taking the turn in a conversation (CEFR 2020: 139). As far as thematic development is concerned, the learners should be able to present ideas in a logical way and they should be able to tell a story or develop a text or an argument (CEFR 2020: 139). The fourth scale is coherence and cohesion, which “refer[s] to the way in which the separate elements of a text are interwoven” especially on a sentence level (CEFR 2020: 140). Learners should for example be able to link elements and use connectors (or discourse markers) and vary the use of cohesive devices (CEFR 2020: 141).

Looking at these key concepts, it becomes clear that aspects of the scales are related with the functions of the discourse markers *so* and *okay* that had been identified in the analysis (see section 4.1.4.). Table 44 depicts an overview of selected scales in the CEFR and identified functions in the conducted analysis.

TABLE 44: SELECTED SCALES AND RESPECTIVE DM FUNCTIONS

Selected key concepts (CEFR)	Functions of <i>so</i> (1-7) and <i>okay</i> (8-15) in TaCoCASE
Turntaking	(2) action prompt, (6) hesitation marker and floor-holding device, (8) Pre-close closing marker , (9) Backchannel, comprehension signal, (11) Agreeing, sealing outcome , (13) Comprehension check, general check, (15) Intersubjectivity
Thematic development	(3) topic developer and sequencer, (12) Topic starter & change , linking device, (10) Return to topic
Coherence & Cohesion	(1) result and inference, (4) connectionmarker, (5) marker of emergence from incipency, (14) Displaying pragmatic competence & fluency

The learners could either be provided with the list (table 44) and the different functions or they would have to identify different functions from the different examples (see section 4.1.4.) themselves. Another possibility is to provide a list of functions and examples and the learners would have to match the functions to the respective example so that they identify the functions and the different uses of the discourse markers.

Some possible teaching tasks for teaching discourse markers based on the findings in section 4.1.4, will be presented in the following.

(3) Teaching Task: Discourse markers

Task 1: Listen to (read) the following conversation and write down all discourse markers you can hear (identify). (03SB149BO03 ; #38:00#)

Task 2: Describe the discourse setting of the example and identify different functions of the discourse marker *so*. Compare your results with your partner and explain the functions in the context of the conversation to each other.

03SB149BO03 (#38:00#)

- BO03: [yeah, yeah maybe if it's your-
.. when your speaking English,
that's is not your native language
yeah that would be hard.
(4.0) so how long is your recording now?
- SB149: uhm I guess about forty minutes.
yeah .. thirty eight.
- BO03: okay.

SB149: so I guess we could call it a day?
 BO03: yeah.
 () anything you want to say?
 SB149: no.
 [I'm- I'm just]
 BO03: [any story]
 SB149: [I'm] just happy .. that I was able to talk to you,
 I really enjoyed the time,
 BO03: yes it was nice meeting you.
 SB149: yeah.
 okay so uhm have a lovely day.
 BO03: thank you ... good luck with your studying.
 SB149: thanks ... same for you.
 BO03: thank you.
 SB149: bye.
 BO03: bye. #00:38:30-5#

In task 1, the learners should identify different discourse markers in the given example. This task could be used in a unit on discourse markers at an intermediate levels after the learners have been familiarized with the concept of discourse markers. They could apply their theoretical knowledge in this example by identifying the different discourse markers (level 2 in Bloom's taxonomy) and before identifying different functions of one particular discourse marker, namely *so* (see task 2). In discussing the results with a partner, the learners have to explain and argue their findings, which on the one hand draws on a high level in Bloom's taxonomy and which on the other hand is a learner-centered communicative approach as the learner has to speak and thus produces output.

Both tasks make the learners aware of the use and functions of discourse markers, which builds on the learners' pragmatic competence and thus on their communicative competence in general.

5.2. TACOCASE AND COMMUNICATIVE LANGUAGE STRATEGIES

This section looks at **communicative strategies** as defined by the CEFR. I will show to what extent examples from TaCoCASE can be used to illustrate different communicative strategies, in particular **interaction** and **mediation strategies**. The examples could be directly used in the classroom in an inductive and learner-centered approach. The learners could work with these examples and identify different strategies that are used by the participants in the conversations. In a next step, the learners could think of additional strategies and then apply their knowledge in simulations in the classroom or in real world intercultural encounters. In this transfer task, the learners apply different interaction and mediation strategies themselves and learn how to use them. The communicative strategies are not only presented as input and in receptive activities, but the learners should also produce output, which means utterances and speech while using these strategies. Not only input and output are concerned as second language acquisition theories (see Krashen & Terrell 1988; and Swain & Lapkin 1995), it also draws on the third language acquisition theory, namely the interaction approach (Long 1996). If the learners use the communication strategies themselves in interaction, they negotiate meaning – in a linguistic sense – and notice if the communicative act they wanted to fulfill was successful or not (Long 1996). They adapt the strategies and come up with a set of strategies they are able to use themselves in interaction with interlocutors. Noticing already happens in receptive activities when the learners try to find and understand the communicative strategies the speakers in the corpus used.

Based on the theoretical framework of communicative strategies and on the requirements for EFL teaching, the following objectives should be reached by the learners:

The objectives for communicative language strategies are:

1. The learners know different communicative strategies, in particular interaction and mediation strategies.
2. The learners can apply interaction and mediation strategies in different situations.
3. The learners evaluate their use of different communicative strategies and come up with a set of strategies that they find useful for themselves.

5.2.1. INTERACTION STRATEGIES

Interaction strategies are communicative strategies that speakers use to interact with another speaker in a conversation. Communicative strategies in SLA research are closely related to the **interaction approach**, established by Michael Long (1996), which is one of the main second language acquisition theories. The central aspect of the interaction approach is the negotiation of meaning, which Michael Long described as follows.

Negotiation for meaning is the process in which, in an effort to communicate, learners and competent speakers provide and interpret signals of their own and the interlocutor's perceived comprehension, thus provoking adjustments to linguistic form, conversational structure, message content, or all three, until acceptable level of understanding is achieved (Long 1996: 418; emphasis in original)

Long focused on conversations between NNS and NS and how the speakers negotiate meaning in interaction. Successful communication and the completion of the communicative goal is central to the speakers, which simply put means that in negotiating meaning, the speakers "[...] make what they say comprehensible to their interlocutors" (Long 1996: 418). He found several devices or strategies that the speakers applied (see Long 1996: 418):

- Repetitions
- Confirmations
- Reformulations
- Comprehension checks
- Confirmation checks
- Clarification requests

Long found that these strategies were used for two reasons: (1) "strategically, to avoid conversational trouble" and (2) "tactically, to repair communication breakdowns when they occur" (Long 1996: 418).

Another approach to communicative strategies is an interactional, sociolinguistic approach, which also takes social and cultural factors into account. Following this approach, "diversity of background and communicative conventions come to take on important signaling functions in everyday interaction" (Gumperz 1982: 7). Discourse strategies are used in interaction to facilitate the process of joint meaning making. Marie-Louise Brunner (2021) and Monica Antonello (2022) looked at communication

strategies in the sister corpus of TaCoCASE, ViMELF, and thus focused on communication strategies in ELF. They both followed an interactional sociolinguistic approach. Antonello found the following verbal communication strategies in ViMELF (see Antonello 2022: 202 ff):

- Backchannels
- Lexical anticipation
- Lexical suggestion and correction
- Overt multilingual resources
- Reformulation
- Repetition
- Spelling

Brunner (2021) investigated communication strategies and negotiation of meaning in ELF and particular in ViMELF and came up with a taxonomy for innovative language use, which includes the following elements (see Brunner 2021: 130 ff):

- Innovative language use: plurilingual resources, approximation, *ad hoc* innovation
- Code-Switching
- Definitions
- Non-verbal behavior

It is not very easy to compare communication strategies across both approaches – the interaction approach in SLA and the interactional sociolinguistic approach – as the terms are usually very different. When looking at the lists of strategies mentioned above and comparing them, it becomes clear that the underlying concepts or strategies are quite similar. Additional and new strategies that were found in ViMELF, which were not taken into account in the interaction approach, are spelling, using overt multilingual resources (or code-switches) and non-verbal behavior. Backchanneling is one way to signal comprehension and giving feedback to the listener and co-operating (Bjørge, 2010; Cogo & Dewey, 2012, Schegloff 1982), which could be counted as confirmation. Lexical anticipation, lexical suggestion and correction (Antonello 2022) as well as innovative language use (Brunner 2021) and definitions (Brunner 2021) are not explicitly mentioned in the interaction approach, even though they could probably be counted as reformulation in a wider sense. The strategies in the interaction approach are

more communicative strategies in the sense that misunderstanding or non-understanding was due to lack of language or due to linguistic reasons. The strategies found in an ELF context expanded the scope and there are two possible reasons for misunderstandings: linguistic as well as intercultural mis- or non-understandings. Some of the strategies, like for example using plurilingual resources or code-switches, are mentioned in the CEFR as part of building plurilingual and pluricultural competence (see section 2.1.5.2.4. and 5.2.2.).

The strategies that are mentioned in the CEFR combine aspect from both approaches of communication strategies and focus mainly on discourse strategies. The CEFR describes descriptor scales for interaction strategies in the CEFR.

Descriptor scales for interaction strategies in the CEFR (CEFR 2020: 87)

- Taking the floor (turntaking)
- Co-operating
- Asking for clarification

The key concepts operationalized for these scales include the main strategies mentioned in the interaction approach (Long 1996). These key concepts for interaction strategies as described by the CEFR are presented in the following figure (figure 55).

Key concepts operationalized for interaction strategies (CEFR 2020: 87):

Taking the floor*	Co-operating	Asking for clarification
• initiating, maintaining and ending conversation • intervening in an existing conversation or discussion, often using a prefabricated expression to do so • gain time to think	• confirming comprehension (lower levels); • ability to give feedback and relate one's own contribution to that of previous speakers/signers (higher levels); • summarising the point reached in the discussion in order to take stock (B levels); • inviting others to contribute	• indicating comprehension or a comprehension problem (lower levels); • requesting repetition; • asking follow-up questions to check comprehension or request more details

FIGURE 55: KEY CONCEPTS OPERATIONALIZED IN THE CEFR FOR INTERACTION STRATEGIES (CEFR 2020: 87)

*turntaking: also pragmatic competence (discourse competence)

Taking the floor or turntaking has already been discussed as aspect of pragmatic competence (or discourse competence) and it is repeated again at this point (and in the CEFR) as it is also an interaction strategy. The key concepts that are listed in the CEFR are in fact exactly the same. Co-operating means adapting “cognitive strategies” such as “framing, planning and organising ideational content” as well as “collaborative strategies” such as “handling interpersonal, relational aspects” (CEFR 2020: 87). The **co-operating** scale will be further developed in the section on mediation strategies in terms of “facilitating collaborative interaction with peers” and “collaborating to construct meaning” (CEFR 2020: 88). In general, this scale is concerned with “discourse moves” in order to develop a discussion (CEFR 2020: 88). **Asking for clarification** is a strategy that is used in order to “intervening in an interaction to indicate whether one is following or not” (CEFR 2020: 89), which includes “indicating comprehension or a comprehension problem”, “requesting repetition” or “asking follow-up questions to check comprehension” (CEFR 2020: 89).

The new strategies found in ELF research and in ViMELF (see Brunner 2021 and Antonello 2022), like backchannelling and spelling are strategies that I also found in TaCoCASE (see table 43). Some other strategies that Antonello and Brunner investigated and found in ViMELF, like for example code-switching, innovative language use, lexical anticipation, lexical suggestion and correction, are strategies that could count as mediation strategies when mediating a new concept and should be added in the CEFR as communication strategies and in particular as mediation strategies.

In order to match examples from the corpus with selected aspects from the descriptor scales, I will present and explain the corpus examples and the respective strategies used in the following paragraph in greater detail. These examples could be used in a classroom setting, which means that the learners could retrieve the strategies from the examples. The learners could then use and adapt these strategies in new situations and come up with a selection of strategies they can apply themselves.

Selected examples taken from TaCoCASE

Example (21) draws on the scale of **taking the floor** (or turntaking) and is an example for the key concept of **gaining time to think**.

(21) Taking the floor: gaining time to think – holidays (06SB157BO11; #42:00#)

- SB157: maybe in terms of holidays we are just re-,
.. yeah .. mo- most of them are just religious and,
.. H uhm.. free of wo:rk and that's it, ((chuckles))
- BO11: {nods}
(1.2) yeah .. (definitely) yeah,
(1.6) **I'm trying to think**, {puts hand against chin}
uh like .. u:hm .. there is >Veteran's Day< too,
which is .. u:h,
.. like just honouring people in the military,
and uhm >**I'm trying to think of<**,
other holidays that they don't .. have other places,
{looks to side} #00:42:32-2#
(1.2) uhm ... yeah
Columbus Day obviously,
... Thanksgiving .. Fourth of July,
but yeah,
I mean those are >the other ones I can think of<,
out of the top of my head,
"OH" Labour Day, {raises eyebrows}
maybe they still don't understand why ((laughing)),
it's called (Labour) Day,
- SB157: what? {moves closer to screen}
BO11: oh it's called Labour Day?
SB157: oh Labour Day yeah,
(I've heard of it) yeah

In example 21, BO11 is taking, or in fact, holding the floor by **gaining time to think**. The participants are talking about holidays and special holidays in their respective country. SB157 implies that she cannot think of more German holidays as they were talking about a few already. After quite a long pause (1.2 seconds), the American speaker takes the floor in order to prevent a communication breakdown. She first agrees by saying *yeah* and nodding at the same time ({nods}) and then only says out loud what she is thinking (*I'm trying to think...*), followed by a thinking gesture ({*puts hands against chin*}). She is thus not contributing to the conversation content-wise, but is holding the floor in order to prevent silence in the conversation. She uses many hesitation markers like *uh*, *like...* *uh:m*, accompanied by pauses to hold the floor and to mark her uncertainty and hesitation. At the end, she

repeats the phrase *I'm trying to think* and uses again a thinking gesture (*looks to the side*) to support her uncertainty and hesitation.

The following example (22) is concerned with the scale of **co-operating** and the key concept of **confirming comprehension**.

(22) Co-operating: confirming comprehension (04BO16BI41; #00:30#)

BO16: .. kay:.
(2.1) m:ight be having some problems:.
{keeps unplugging/ replugging headphones}
(1.4) well that's unfortunate?
(1.0) okay.
(1.5) {is searching through a drawer}
I can try: .. any of my other .. headsets.
BI40: **.. could_you .. not HEar .. [my voic:e]?**
BO16: [{changing headsets}]
(1.9) can you hear me?
BI40: **.. I could hear you.**
could you hear me? {pointing with index finger to herself}
BO16: .. okay.
... I cannot hear you:. [{tries replugging headphone again}]
BI40: [o:hh],
BO16: (1.5) how about now?
BI40: uhm-
.. I could hear you. #00:00:59-5#
BO16: ... oh- okay.
this is biza:rre.
fai:led .. play test tone.
this is .. °NOT good°.
.. okay.
give me just a moment please.
I'm so:rry:.
BI40: that's alright.
BO16: (4.2) (°that sucks°.)
AH.
can you hear me?
BI40: .. yeah. {drinks from a bottle}
BO16: **... I can hear you.**
[okay].
BI40: [((ehh))] h. {recaps bottle}
BO16: ... THAT was really bizarre.
.. okay.
.. HELLO:?
BI40: .. hi?
BO16: ... well .. fl:Rst.
.. I wanna know,
cause I'm BAD at this,
.. how do I pronouns your NAME? #00:01:31-0#

Example 22 shows very clearly how the speakers are **co-operating** and how one speaker is **confirming comprehension** by using repetitions and backchannels. The participants have understanding problems due to technical issues. The American speaker (BO16) cannot hear the British speaker (*I cannot hear you*) , while the British speaker (BI40) confirms that she can hear the American speaker (*I could hear you*). He (BO16) keeps plugging and unplugging his headphone (*[[tries replugging headphone again]]*) and keeps asking if the British speaker can understand her and she very patiently cooperates by saying: *I could hear you*, which she repeats a few times after BO16 is still trying to figure out the problem. There is also backchanneling and again an affirmation (*I could hear you*). She signals understanding and confirms comprehension while at the same time showing listenership. It is interesting that the British speaker keeps repeating the exact same phrase over and over again very patiently and thus signals comprehension. After the American speaker presumably fixed the problem and they could actually understand each other, they start the actual conversation.

Example (23) represents the scale of **asking for clarification** and the key concept that is used is a **asking a follow-up question**.

(23) Asking for clarification – Follow-up question (06BO18BI38; #02:00#)

- BI38: O:H you had THANKSgiving,
what was that like,
'cause I don't know what Thanksgiving is? #00:02:00-8#
- BO18: OH okay,
well so the story behind Thanksgiving is u:hm
... .t ... way back whe - do you know the story or am I boring you with that
- BI38: >No no no< I don't know,
all I know is there is a turkey
[...]
- BO18: .. one of the holidays everybody is supposed to get together.
you know your whole family is supposed to come into TOWN,
that sort of thing.
- BI38: **is it bigger than CHRISTmas though?**
or is it kind of the same,
- BO18: uhm I would say it is .. it's on part,
probably slightly <LESS> important
in most people's minds than Christmas
- BI38: oh_okay. #00:03:32-1#

This excerpt (23) has already been discussed (see example 18), but with a different focus. The participants are talking about Thanksgiving, a concept the British speaker (BI38) is trying to grasp and the American speaker (BO18) is explaining the history behind it (see full explanation in example 18). The British speaker is **asking for clarification** about the whole concept of Thanksgiving because she does not know about the concept (*oh, what is it like, 'cause I don't know what Thanksgiving is*). After the very detailed explanation of the American speaker, the British speaker asks: Is it bigger than Christmas though? She signals understanding and asks a follow up question in order to signal interest and to better understand the value of the holiday and tries to compare it to a holiday (or concept) that she knows or can at least relate to. She is clarifying meaning by **asking the follow-up question** and is verifying the concept she has in mind with the concept that her interlocutor, the cultural expert, is providing. By asking the follow-up question, she also signals curiosity and openness for her interlocutor's culture in acknowledging his expert status while at the same time claiming co-expertise to some extent as she knows about some aspects of Christmas as well. She shows that she is interculturally sensitive in verifying the concepts and asking about details and showing interest with the **follow-up questions**, which prompt further explanation by the cultural expert.

Teaching Tasks – Interaction strategies

Task 1: Listen to the conversation (read the example) and match the following interaction strategies to the respective examples: Taking the floor, co-operating and asking for clarification.

Task 2: Group puzzle: Go together in groups and pick one of the example (21-23). How did the participants use interaction strategies, i.e. what did they do or say in order to take the floor, co-operate or ask for clarification? Can you think of more strategies they could have used? Present your findings in class.

Task 3: Work together with a partner. Imagine you are talking to your friend in the U.S. (in a videocall). One of you is the American student and one is the German student. Act out a situation in which you use one of the strategies mentioned above. After the simulation, take some time to reflect: Did you successfully use the interaction strategy? Were they useful? Did you add or combine other strategies.

Task 1 is a matching exercise in which the communicative strategies the CEFR suggests, taking the floor, co-operating and asking for clarification are provided and the learners should find instances of these strategies in the conversation – either in the text or in the audio, depending on the teacher’s preference (focus on reading versus listening skills).

Task 2 is a cooperative learning exercise in which the learners have to work in groups and therefore not only fulfill the task but also apply five elements of cooperative learning, such as positive interdependence, face-to-face interaction, individual accountability, social skills and group processing (see Bonnet et al. 2017: 173; see also Johnson et al. 2008; Bonnet 2009). The learners should find the strategies mentioned in the CEFR and they should think of more strategies that could be useful.

Task 3 is a task-based language learning (TBLL) activity (Meyer et al. 2022: 69; see also Decke-Cornill et al. 2015: 120), which also fulfills the characteristic elements of cooperative learning as mentioned in task 2. In this task, the learners pretend as if they were an American and a German student and they should make up and act out a dialogue using the strategies they learned before. They should also reflect on their use of the strategies, whether the strategies helped them to interact, whether they used the strategies correctly or not, and whether they used other strategies apart from the ones they have already learned.

Task 1 is a beginner level task, which corresponds to levels 3 and 4 (application and analysis) of educational objectives in Bloom’s taxonomy. Task 2 and 3 are intermediate and advanced learners’ tasks, which correspond to higher cognitive levels such as synthesis (level 5) and evaluation (level 6) in Bloom’s taxonomy (see section 2.1.1.).

5.2.2. MEDIATION STRATEGIES

The foundations of **mediation** have already been discussed with regard to establishing intercultural competence (see section 2.1.5.2.). The goal of mediation in EFL teaching can be summarized with the following quote from the CEFR: “In mediation, the user/learner acts as a **social agent** who **creates**

bridges and helps to **construct or convey meaning**" (CEFR 2020: 90; my highlighting). Mediation can take place in a communicative or linguistic setting, e.g. when translating and interpreting, and it can take place in intercultural communication, when (new) cultural concepts are being explained and meaning is negotiated. These two kinds of mediation play an important role in EFL teaching and learning; in particular cultural mediation is of special interest in the scope of this work as it contributes to intercultural competence. **Mediating (new) cultural concepts**, according to the CEFR, is characterized by two main aspects: "constructing and elaborating meaning and [...] facilitating and stimulating conditions that are conducive to such conceptual exchange and development." (CEFR 2020: 91). The learners need strategies which they can apply in order to "clarify meaning and facilitate understanding" (CEFR 2020: 117). The CEFR lists two factors for mediation strategies: (1) strategies to explain a new concept and (2) strategies to simplifying a text. As the focus in this work is on establishing intercultural competence and on mediation cultural concepts, I will only look at strategies to explain a new concept in the following. The CEFR mentions three strategies to explain a new concept.

Strategies to explain a new concept (CEFR 2020: 118):

- Linking to previous knowledge
- Adapting language
- Breaking down complicated information

When explaining a new concept, the learner or mediator can link to previous knowledge of the interlocutor so that the interlocutor can easier understand the new concept. The link to previous knowledge could be established through "posing questions to encourage people to activate prior knowledge", "making comparisons and/or links between new and prior knowledge" and providing examples or definitions (CEFR 2020: 118). Another strategy to explain a new cultural concept is adapting language, which can be described as a shift in language use either in style or register (CEFR 2020: 118). The key concepts as per the CEFR include "paraphrasing", "adapting delivery" and "explaining technical terminology" (CEFR 2020: 118). The third strategy to explain a new concept is "breaking down complicated information" into smaller parts so that the information can be processed

more easily. This means a process could be broken up into a series of smaller steps. In writing, ideas could be presented in bullet points, which are usually formulated clear and very concise, so that they are easier to read and understand. The separate presentation of main points in a chain of arguments also belongs to the key concepts for breaking down complicated information (CEFR 2020: 118).

Key concepts operationalized for explaining a new concept (CEFR 2020: 118)

Linking to previous knowledge	Adapting language	Breaking down complicated information
• posing questions to encourage people to activate prior knowledge; • making comparisons and/or links between new and prior knowledge; • providing examples and definitions.	• paraphrasing; • adapting delivery; • explaining technical terminology.	• breaking a process into a series of steps; • presenting ideas or instructions as bullet points; • presenting separately the main points in a chain of argument.

FIGURE 56: KEY CONCEPTS OPERATIONALIZED IN THE CEFR FOR EXPLAINING A NEW CONCEPT (CEFR 2020: 118)

I will present selected examples from TaCoCASE that match the strategies mentioned in the CEFR in the following.

Selected examples from TaCoCASE

Some examples from TaCoCASE will be used to explain selected strategies from figure 56. The examples could be used in the EFL classroom and learners could retrieve the strategies that were used by the participants and understand how the participants used the mediation strategies. In a next step, they could come up with own strategies and apply some of them in a simulation or a real world (intercultural) situation outside the EFL classroom.

(24) New concept: Linking to previous knowledge: Comparison – Brezel (07SB156BO10; #32:00#)

SB156: [Brezel ((German (0.2))) right]?
Brezel ((German (0.2))).
do you know that?
BO10: mhm. {drinks out of a mug}
SB156: OH MY GOD uhm,
how do I explain it Brezel ((German (0.2)))? #00:32:30#
you have never heard of a **Brezel** ((German (0.2)))?
BO10: **betsin?**
SB156: Brezel ((German (0.2))) **B-R-E-Z-E-L**.
BO10: no .. uh-huh. {shakes head} {types something & looks to screen}
[{typing sound (0.5)}]
SB156: [NO you eat] it in Bavaria and it's,
BO10: OH .. [YE:AH]=
SB156: =[Brezel ((German (0.2)))]
BO10: pretzel.
SB156: Bretzels.
BO10: ((chuckles)) [But you say Brezel ((German (0.2)))]=
SB156: =[() Brezel ((German (0.2)))]
BO10: with a b Brezel ((German (0.2)))?
SB156: b it's actually Brezel ((German (0.2))) yeah Brezel ((German (0.2))) [B]=
BO10: =[yeah] and in English,
we just say **pretzel with a P**. #00:33:00#
SB156: okay.
BO10: yeah.
SB156: actually **it's spelled with B**.
BO10: uh-huh.
SB156: (2.0) and uhm ... you serve it with uhm .. sweet honey? {tilts head}
BO10: mhm. {nods}
SB156: white sausage.
BO10: OOH.
SB156: =[you know] Germans call it Weisswurst ((German (0.3))).
BO10: mhm.
SB156: Weisswurst ((German (0.3)))
BO10: OH >yeah – yeah – yeah< okay.
[it's a the same–]=
SB156: =[what's the word in English]
BO10: I think it's the same word in English,
just butchered,
we say bratwurst.
SB156: .. OH okay. {nods}
BO10: yeah. #00:33:30#

In example 24, we find the German speaker **explaining a new concept**, in this case *Brezel*, to the American speaker. The context suggests that the American speaker has a hard time understanding what the German speaker means at the beginning and the German speaker is trying very eagerly to explain it by saying where it is eaten in order to **link to the American speaker's prior knowledge**. He spells out the word B-R-E-Z-E-L and the American speaker is typing something – she is probably searching for the word online. She says *OH YEAH*, using the interjection, indicating that she just found

out what the German speaker meant and that she knows what a *Brezel* is. It can be suggested that she might have seen a picture online, which she **compares to her concept of a pretzel**. The interlocutors have a little argument about the correct spelling and the differences in American English and German. The German speaker again tries to link to previous knowledge of the American speaker by saying that *you eat it with sweet honey and Weisswurst* (he uses a code-switch for white sausage again to be very precise about the concept). They compare the German *Brezel* and *Weisswurst* and the American *pretzel* and *bratwurst* in order to clarify meaning and show cultural expertise for the respective version of the concept. The German speaker is claiming cultural expertise in fact very explicitly by informing the American speaker about the correct spelling (*actually it's spelled with B*) and how it is eaten. They both come to the conclusion that their concepts are not very different by repeating that it is the same (it's a the same) to signal confirmation and understanding.

(25) New concept: Adapting language: Paraphrasing and explaining terminology – Spanish medical interpreting class

BO06: it's kind of interesting?
and the:n...e:r,
..let's see.
I'm taking a **Spanish medical interpreting class**. #00:13:30-0#
(3.) and [that],
SB152: [what's that].
BO06: what's that?
SB152: yeah what is that.
BO06: uhm in it-,
it's basically to **prepare you for the test**.
e:hm so it's like a course a:ll semester long,
...and then after the course is over,
you can go and take like the...**certification test**.
(2.0) ehm in America **we don't have like a national**
...well we have like a **national exam**,
but that is **for like a master level interpreter**. #00:14:00-0#
but **if you just going to work in one state**,
at one **hospital** or two hospitals,
...then it's just up to local people to decide
(4.0) [so yeah].
SB152: [okay].

In example 25, we can see how a new concept is explained by adapting language and in particular by paraphrasing and explaining complicated terminology. The American speaker is talking about a class he is taking – a *Spanish medical interpreting class* – which is quite a technical term. The German

speaker does not understand the concept, either because she does not know what an interpreting class is or because the concept or the system is different in Germany and asks a question to clarify meaning. The American speaker reacts to the question and interprets it as a prompt to explain the concept. He is engaging in an explanation by **adapting his language** and **paraphrasing the concept**: *it's to prepare you for the test, it's all semester long, it's a certification test*. By providing this detailed information or explanation, the American speaker is describing elements and characteristics of the test in order to support the German speaker in her understanding process. It could be suggested from the context that the American speaker tries to provide information so that the German speaker draws on similar concepts she know and recalls it by hearing these characteristics. He is thus describing the new concept by paraphrasing and providing more information in order to clarify meaning. After the explanation, the German speaker provides feedback and signals understanding and comprehension in the end by using *okay* as a backchannel.

(26) New concept: Breaking down complicated information – Sautéed mushrooms (07SB156BO10, #26:00#

- SB156: what kind of burger?
like cheese burger or [just]=
- BO10: =[u:hm] .. my favourite would be,
it's called like a **patty melt**.
it's a burger on .. pieces of bread,
instead of a bun, {imitates putting together a burger}
(2.3) and uhm .. my favourite is to get one of those,
with like sautéed mushrooms on it?
and Swiss cheese.
- SB156: salted mushrooms?
- BO10: sautéed mushrooms. {raises eyebrows}
- SB156: °oh° okay. {puts head down}
- BO10: (2.1) do you know **sautéed**?
... **like how you cook them**?
- SB156: sautéed?
- BO10: **here I'll – I can send it to you.**
{types something & looks down}
{**typing sound** (0.3)}
{Skype message sound}
- SB156: (5.0) °oh° okay yeah go ahead.
- BO10: ... I actually think it's a French word sauté.
... It's **like to put in a pan**,
and like .. **you put in some oil or butter**,
and you just .. **get them warm**, {makes rolling gesture}
and a little uh brown.
- SB156: <sauté> {types & looks to screen and down}
{typing sound (2.8)}

>OH< okay.
BO10: yeah .. so some **sautéed mushrooms**,
with Swiss cheese on a burger,
that's probably my <absolute favourite>. {smiles}

In example 26, the American speaker (BO10) talks about her favorite burger and introduces a new concept – *a patty melt with sautéed mushrooms*– that she likes to eat. SB156 is asking for clarification and signaling non-understanding by repeating the (wrong) word *salted mushrooms?* with raising voice, indicated by the question mark in the transcript. He prompts the interlocutor to repeat the word to clarify meaning and pronunciation of the word. BO10 notices that SB156 did not understand what she said (he used the word *salted* instead of *sautéed*) and asks if the German speaker knows what *sautéed* means (*do you know sautéed? like how you cook them?*) by repeating the word and asking about the process how to cook the mushrooms. BO10 is explaining the new concept by using multimodal features (she sends a picture in the chat) and by **breaking down the process** of making sautéed mushrooms into a series of smaller steps: *like to put in a pan, you put in some oil or butter, you just get them warm and a little brown*. SB156 signals understanding by saying >OH< okay. He uses >OH< as an interjection, which he speaks with a louder voice (indicated by the capital letters in the transcript), in combination with *okay* as a marker of comprehension. It becomes clear that by using the interjection in combination with the comprehension marker, he just found out about the meaning and now knows what *sautéed* means and thus what the concept means. BO10 repeats the whole concept – *sautéed mushrooms with Swiss cheese on a burger* – and repeats that it is her favorite burger as an affirmation from what she stated at the beginning and in order to emphasize the importance to her.

Another example that could be used in this category – breaking down complicated information – is example 19 (Sauerkraut).

Implications for teaching

The following teaching tasks use example 26 as text or conversation (video), which should be provided together with the tasks. Since the term *sautéed* is not very common and the language that is used in the example is not very easy, I would suggest to use this example and the tasks at an intermediate level.

Teaching Tasks

Task 1: Did you understand what *sautéed* means? Find a picture of a burger with sautéed mushrooms online and check, whether the description of BO15 was precise.

Task 2: Did you understand what *sautéed* means? Explain to your neighbor in your own words what *sautéed* means.

Task 3: Imagine you are visiting a friend who lives abroad and you both talk about your favorite food. Explain to your friend what your favorite (traditional German) food is and be very precise so that he/she knows what you are talking about.

In **task 1**, the learners apply their knowledge and should try to understand the concept (*sautéed*) and provide proof that they understood the concept by finding a picture online of *sautéed mushrooms* and *a patty melt*. They are combining a visual picture of sautéed mushrooms with the term and the concept that they just learned, which is a multimodal way of understanding and acquiring a new concept.

Task 2 is an explanation task which has a communicative element. The learners have to explain in their own words what the concept means, which needs several strategies to explain the concept, such as paraphrasing or breaking down the process, for example.

Task 3 is a task-based language learning (TBLL) task, in which the learners need to adapt the strategies they learned and use it in a different context. The learners pretend as if they were visiting a friend abroad and they are talking about their favorite food from their home country. This represents a very authentic situation which the learners could come across at some point in their life and they should be able to act communicatively and interculturally in this situation. The learners could use the strategies that they learned before and adapt the strategies by explaining their favorite food. It is also a very

motivating task for the learners as talking about their favorite food is a topic that is very close to their real world and also relevant for learners to share this piece of personal information. Task 1 and 2 correspond to the comprehension and application level (level 2-3) in Bloom's taxonomy, while task 3 draws on a higher cognitive level, which is evaluation (level 6) in the taxonomy.

Further mediation strategies and discourse structure

The analysis in section 4.2.2. suggests that there are more mediation strategies to explain a new cultural concept than mentioned in the CEFR. The mediation strategies that the speakers in TaCoCASE used to explain a concept known to both speakers is that they **shared experience** about the concept and drew comparisons between the respective concepts that they know. If a concept was only clear to one speaker, the speaker whose culture the concept was from, explained the concept to the other speaker. If the speaker wanted to talk about a concept from the interlocutor's culture and was not sure about the details or if his/her concept was right, the speaker usually tried to verify the concept or asked for clarification (see figure 51). Examples were discussed in section 4.4.2. and these examples (examples 14 – 18) could as well be used in a classroom to teach the strategies that the speakers in TaCoCASE applied.

The speakers in TaCoCASE made use of most of the strategies mentioned by the CEFR, however there were also **new strategies** that the CEFR does not mention, for example spelling, using code-switches, using non-verbal elements like gestures to explain a food item or multimodal features like sending a picture in the chat (see table 42). The CEFR focuses very much on mediation on a verbal level for explaining new concepts and disregards non-verbal elements in this category, which I would suggest to add for the next version of the CEFR. Code-switching is mentioned as part of plurilingual resources in the CEFR with the key concept of "blending and alternating between languages where necessary" (CEFR 2020: 127) and spelling probably could count as "adjusting language according to the linguistic skills of interlocutors", even though spelling is not explicitly mentioned as communicative strategy in

the CEFR. Spelling and code-switching are communicative strategies that were especially found in ELF research (see Kirkpatrick 2007 for spelling and Brunner and Diemer 2018, Brunner 2021, Cogo and Dewey 2012 for code-switching). Brunner (2021) and Antonello (2022) also identified these strategies in ViMELF, the sister corpus of TaCoCASE, which underpin the importance of these strategies in interaction. The new strategies that Brunner (2021) and Antonello (2022) found in ViMELF and that I partially also identified in TaCoCASE (see table 41) should be added as communication strategies and in particular as mediation strategies when explaining new concepts.

A major part of mediating concepts has to do with explaining a new concept if the interlocutor does not know the concept (at first). In section 4.2.2., I looked more closely on how the speakers in TaCoCASE explained new concepts in terms of discourse structure. I came up with a general discourse structure that can be used when explaining a new concept that consists of the following five steps (see also table 42):

- Introduction of new concept or topic
- Check level of knowledge
- Evaluation
- Explanation or elaboration
- Result

In a task, in which the students should explain (new) cultural concepts, they could follow this discourse structure and apply some sample actions and expressions that they could easily use, for example:

- *Do you know ... ?* (when introducing a new concept)
- *Oh, you should try it* (when evaluating the concept)
- or the interlocutor could say *oh that's interesting* as a result after the explanation.

More examples of expressions that the learners could use are depicted in table 42. The learners could identify the strategies and the discourse structure by looking at the examples presented in this section or in section 4.2. They could extract the strategies, look at how they were used and make use of the strategies themselves, which would be a bottom-up and inductive approach. This example will be further discussed as a case study (see section 5.2.2.). In a top-down or deductive approach, the

teachers could provide the discourse structure and the learners would have to find these elements in the examples.

5.3. TaCoCASE AND INTERCULTURAL COMPETENCE

TaCoCASE consist of conversations between students with different language and cultural backgrounds, which classifies the conversations as **intercultural communication** or **intercultural encounters**. Intercultural competence usually unfolds in intercultural situations and in intercultural communication. Section 2.1.5.2. already looked at intercultural competence from a scientific and didactic point of view. The criteria for intercultural competence – according to the CEFR and the German education standards – were described in detail and a synopsis with regard to intercultural aspects in the German federal state curricula was conducted. Having these definitions, criteria and objectives in mind, the focus in this section is on establishing intercultural competence in the EFL classroom and the elements depicted in the CEFR – similar to the other sections in the application part. Examples from TaCoCASE can be used to identify strategies in order to act interculturally sensitive and competent. The learners could then apply these strategies and make use of the strategies themselves in simulations (as part of TBLL) or in real-life intercultural encounters.

As mentioned in section 2.1.5.2., intercultural competence is usually not taught directly and only in combination with communicative competence. It is usually also not assessed in tests or even in final exams (e.g. Abitur), even though its significance, especially with regard to future skills, globalization and migration, is acknowledged. One exception is the **DESI-IKK assessment test**, which is one of the larger studies that investigated intercultural competence as part of the DESI study (“Deutsch Englisch Schülerleistungen International”) among 9th-graders in Germany. The task for the learners was to predict the probability of a possible reaction, which was described in an answer item, and thus answer questions on the basis of a critical incident that was presented to them. The questions were: What happened? How did the participants feel? How would you react? What can you learn from this story?

(see Hesse & Göbel 2007: 265). The answer items were based on Bennett's (1993) Development Model of Intercultural Sensitivity (DMIS) with the six steps of intercultural sensitivity (from ethnocentrism to ethnorelativeness) (Bennett 1993: 182; see also Göbel 2011; section 2.1.5.2.). Using critical incidents in teaching and assessment seems to be a good method to establish intercultural competence. The examples from TaCoCASE could serve as critical incidents – or without a negative connotation – as positive intercultural examples for successful intercultural communication. Learners could understand and extract strategies for intercultural communication and use them in simulations in the classroom or own intercultural encounters outside the classroom.

The **topics of the intercultural analysis** – food, customs and traditions and popular culture – are topics that the learners can relate to and depict aspects of their daily life. According to the JIM-study, for example, TV series play an important role in young adults' lives (12-19 years old) as they watch TV series most frequently (girls: 22%; boys 20%), followed by crime and mystery series, sitcoms and daily soap operas (JIM 2023: 37). The study also found that *Youtube* (63%) and *Netflix* (50%) are frequently used by young adults (JIM 2023: 39). The participants in TaCoCASE also mention that they use these media channels or streaming services, which supports the findings in the JIM study. The fact that the participants in TaCoCASE are roughly the same age as the learners make the conversations very comprehensible and relatable for learners. Thus, the intercultural aspects are very interesting for the students and could increase the learners' motivation and investment in the EFL classroom (Deci & Ryan 1993; Darvin & Norton 2015).

In order to establish intercultural competence, examples from these intercultural encounters in TaCoCASE could be used in the EFL classroom. As far as this section is concerned, the objectives for establishing intercultural competence, which should be reached, can be described as follows:

The objectives for intercultural competence are:

- 1) The learners know how to behave in an intercultural competent way in intercultural situations and can name factors that contribute to mutual intercultural understanding.
- 2) The learners can apply different factors of intercultural competence and act in an intercultural sensitive way.
- 3) The learners can critically evaluate their intercultural behavior in intercultural simulations.

Before looking at how these objectives can be reached and how the examples from the intercultural encounters in TaCoCASE can help reaching these objectives, I will take a step back and recall what the CEFR suggests in terms of key concepts for intercultural competence.

(Inter)cultural competence and the CEFR

The concept of intercultural competence has been exhaustively discussed in section 2.1.5.2. with special regard to the CEFR (see section 2.1.5.2.4.). It can be stated that intercultural competence has to do with mediating (cultural) concepts and meaning. Therefore, mediating (new) concepts, as described in section 5.2.2. as a communicative language strategy, also contributes to promoting intercultural competence. The CEFR usually refers to **plurilingual and pluricultural competence** that should be established by the learners, which consists of “building on pluricultural repertoire” as well as “plurilingual comprehension” and “building on plurilingual repertoire” (CEFR 2020: 123). As the focus in this section is on cultural aspects, the communicative or plurilingual aspects will not be taken into account and I will only look at “building on pluricultural repertoire” (CEFR 2020: 123). The different terminology with regard to intercultural competence has already been discussed in section 2.1.5.2.2. and the pluralistic approach and building on pluricultural repertoire as demanded by the CEFR is very closely related to the concept of intercultural competence that underlies this work (see section 2.1.5.2.). According to the CEFR, intercultural or pluricultural competence unfolds when the learners build on pluricultural repertoire (CEFR 2020: 123). The key concepts that the CEFR describes as part of pluricultural competence are the following.

Key concepts operationalized (CEFR 2020: 124):

- “recognising and acting on **cultural, socio-pragmatic and socio-linguistic conventions/cues**;
- recognising and interpreting **similarities and differences in perspectives, practices, events**;
- **evaluating** neutrally and critically.”

These aspects contribute to building pluricultural repertoire and thus to plurilingual and pluricultural competence. As far as **building pluricultural repertoire** is concerned, the CEFR lists the following criteria as descriptors for this scale:

Pluricultural repertoire
• the need to deal with ambiguity when faced with cultural diversity, adjusting reactions, modifying language, etc. • the need for understanding that different cultures may have different practices and norms, and that actions may be perceived differently by people belonging to other cultures; • the need to take into consideration differences in behaviours (including gestures, tones and attitudes), discussing over-generalisations and stereotypes; • the need to recognise similarities and use them as a basis to improve communication; • the will to show sensitivity to differences; • readiness to offer and ask for clarification: anticipating possible risks of misunderstanding.” (CEFR 2020: 124)

FIGURE 57: CRITERIA FOR BUILDING PLURICULTURAL REPERTOIRE (CEFR 2020: 124)

The criteria mentioned in the CEFR should help to negotiate meaning in interaction, create common ground and establish mutual understanding. With the help of the criteria mentioned above (see figure 57), the learners should be able to draw on the key concepts for plurilingual and pluricultural competence and act interculturally sensitive and competent in intercultural situations. These criteria are based on **Byram’s ICC model** and the five skills that should be established: (1) knowledge, (2) skills to interpret and relate, (3) skills to discover and or interact, (4) attitudes and education (political education and critical cultural awareness) (Byram 1997: 34; see also section 2.1.5.2.3.). It also draws on the model of **intercultural sensitivity** (Bennett 1993: 182). Since these criteria are based on established models in EFL teaching (Byram 2021; Bennett 1993) and since the criteria were only added in the 2020 version (2018 version as well) of the CEFR, they are quite up to date (see also section 2.1.5.2.4.). However, it is clear that further descriptors and material is needed in order to implement

them into the EFL classroom and to establish intercultural competence. Throughout this work, I try to close this gap by providing teaching material or examples that could be used in the EFL classroom. The following examples represent situations in which the speakers in TaCoCASE used the strategies mentioned by the CEFR to build pluricultural repertoire and thus establishing pluricultural competence. The examples could be used in the EFL classroom and the learners could retrieve the strategies and apply the strategies themselves in intercultural encounters.

Selected examples from TaCoCASE

(27) Dealing with ambiguity: adjusting reactions and modifying language – octopus (03SB149BO03; #21:00#)

SB149: okay,
 what was the most exotic thing .. you've ate- eaten so far,
 BO03: uh .. uhm .. I was in Italy
 .. this summer and they served us-
 we went on a day long tour,
 and so they served us lunch.
 and .. there was all kinds of seafood.
 SB149: [okay].
 BO03: [that I] wouldn't have normally tried.
 there was like octopus ... and I can't remember all the other ones.
 because I think octopus was the gross- the weirdest one. #00:21:30-4#
 it wasn't gross actually.
 ((chuckles))
 but you could see the like suckers [or],
 SB149: [yeah you can- yeah] you can actually see them yeah,
 BO03: yeah it was purple.
 but i ate it .. it was actually good,
 SB149: yeah .. they taste quite good.
 even though they look quite ...
 BO03: yeah. ((laughs))
 ((chuckles))
 BO03: like just close your eyes .. put it in your mouth. ((laughs))

In example (27), the participants are talking about exotic food they have tried as SB149 asks BO03 what the *most exotic food* was she has eaten so far. BO03 starts telling a story: When she visited Italy and they were served seafood, octopus was the weirdest thing she was served or that she tried. Before using the word *weird*, she starts off with *gross-*, but reformulates and self-corrects her by saying *was the weirdest one*. She **adjusts her reaction** to the situation and **modifies her language** to softens the judgement. even though it is still very negative, she probably adjusted her language and reaction either

(1) because she did not want to insult the Italian cuisine by saying that food that she did not like is *gross* or (2) because she thought *weird* was a more suitable or appropriate word than *gross* as far as her memory of the taste is concerned. The context suggests that it was probably the first reason because a little bit later in the conversation, she even says that *it wasn't gross actually* and even a little bit later she brings up the taste again and says that *it was actually good*. SB149 is keeping his reactions quite to himself, only chuckling a little bit as backchannel to signal active listenership, and asking for clarification (*could you see the suckers*) in order to prevent him from judging the food. Both agree that even though they find the look of octopus not very appealing, they like the taste of it. There is lots of laughing and chuckling involved to create rapport while talking about a delicate topic. At the end, they both make fun of the situation and BO03 shares her survival strategy (*just close your eyes and put it in your mouth*). BO03 is acting in a very culturally sensitive way in this situation and is eating the food in order not to insult the local cuisine and the chefs and she shows cultural openness by trying out regional food.

(28) Different cultures, different norms – bus stop (03SB149BI31; #25:00#)

- SB149: [they-] they just- they just,
 ... {lifts hand} went to the door,
 pushed the button,
 the door slammed open {makes brushing-away gesture},
 and they went out. {lifts shoulders}
 the bus driver did- didn't even,
 .. say a thing {moves hand in circle},
 he just closed the door again and {moves hand quickly} waited. #00:25:01-3#
h and **for us in Germany**,
 that's.. absolutely, that's-
- BI31: [[[chuckles]]]
- SB149: [you- you-] you can't imagine someone {opens hand} .h,
 EVEN if it's CLOSE to your bussto[p] **you would never touch the button,**
because it's red and you have to pay fees,
 [...] {shakes index finger}
 because it's- **it's <emergency use only>.**
 [...]
- BI31: Mh. {shifts position} (2.0)
we just go. ((laughs))

Example 28 shows how **different cultures** have **different norms**. The German speaker (SB149) experienced a situation on a bus in London that would not be possible in Germany: Passengers on a bus in London pressed the emergency exit button, because the bus did not move due to traffic jam.

SB149 tells this story and expresses his disbelief or his surprise about the reaction by the people on the bus because this would not be possible in Germany (*you would never touch the button*) and he was not used to a situation like this. This difference in behavior or in norms was apparently a very defining situation for him as he seems to be very excited and tells the story very lively, using gestures ({moves hand quickly}) and emphasizes some particular words (*EVEN if it's CLOSE*). SB149 is drawing conclusions from this one situation to all British passengers using the bus, which could be interpreted as an overgeneralization here. BI31 is supporting his hypothesis by supporting the behavior of the British passengers (*we just go*) identifying with the culture and claiming expert status as she might have done it before herself. It is interesting to see that they both position themselves with regard to their respective culture: *for us in Germany, this would not be possible* versus *we* [the Brits] *just go*, without ending up in a controversial situation. The participants do not evaluate the situation; they simply describe the differences – which depicts intercultural sensitivity and interculturally competent behavior.

(29) Differences in behaviors: Discussing stereotypes – Austrian cuisine (#31:00#)

- BO10: [...] .. I haven't had much German cuisine. #00:31:00#
 so I don't know how it compares.
 .t ALTHOUGH {looks to side & opens eyes widely} **I was in,**
I don't want to insult you,
 but ((heh)) I was in .. **Austria** uh this past.
- SB156: [okay].
- BO10: [summer] and ... like I don't know,
 if the food is similar at all.
 I know the language.
 is a little bit different,
 the accent is different [so]=
- SB156: =[°yeah°]
- BO10: and I don't know **if there is any contention**, {touches chest with both hands & smiles}
between Austria and Germany <so>. {moves around in chair & smiles}
- SB156: (2.0) so what – what was it like the food? #00:31:30#
- BO10: uhm {looks up}
 .. actually I don't know,
 if I had **anything really traditional**.
 it was just uhm **heavy meats**. {lifts shoulders}
 there was a lot of meat. {frowns}
 so, {lifts shoulders & tilts head}
 that's ((heh)) that's all I can really say,
 about °what it was like°.
- SB156: I don't know I don't know,

about the Austrian cuisine actually.
BO10: mhm. {nods}
SB156: to be honest I don't know about the German cuisine.
BO10: ((chuckles))
SB156: I don't know if we have,
like any **really traditional .. FOOD.**
BO10: mhm.
SB156: it's for a sausage and – and – and,
some kind of potato variations. {moves around in chair}
BO10: mhm.
SB156: potatoes dishes or – or – or
uhm.. what else? {looks up}
sausage sauerkraut ((German (0.3))) °I mean°,
what else are we famous for uh?
BO10: **beer maybe?** ((laughs))
you drink a lot of beer.

In this example (29), BO10 shows a large degree of cultural sensitivity at the beginning of the conversation. When asked about German food, she states that she has not had German food and can only relate to Austrian food (*I haven't had much German cuisine*) and makes clear that she is not sure to what extent Germany and Austria are similar in terms of food. She shifts the scope of the conversation to a personal level by saying *I don't want to insult you*, when bringing up that she was in Austria and had Austrian cuisine when asked about German cuisine originally. It seems that she is very careful in comparing Germany and Austria and is trying to avoid to say something insulting to her German interlocutor and she even states that she is not sure *if there is any contention between Austria and Germany*. She draws on her prior knowledge (socio-cultural orientation knowledge) and describes that she is aware of differences in terms of language and accent in particular between Germany and Austria. At the beginning she is mentioning all these presuppositions and it seems that she wants to make clear that she is very culturally sensitive and does not **want to use stereotypes** or overgeneralizations as she emphasizes that she can only speak from her experience (*actually I don't know if I had anything really traditional*). She then makes the first description or classification of the food and states that it involved *a lot of meat*. The German speaker also admits that he does not know about the Austrian cuisine (*I don't know about the Austrian cuisine actually*) and tries to create common ground by stating that he is also not an expert on German cuisine (*to be honest I don't know about the German cuisine*). They both explicitly state that they are non-experts of the German cuisine

even though it could be assumed that the German speaker at least is a cultural expert in terms of German cuisine. He thus puts himself on the same level as the American speaker and diminishes his “supposedly” expert status due to his nationality. The German speaker (SB156) can also not come up with an example for traditional German food (*I don’t know if we have like any really traditional FOOD*). SB156 names some food items that he thinks are traditional, like *sausage, potato, sauerkraut*. BO10 then is suggesting *beer maybe*, and with using *maybe* and a rising voice (indicated by ? in the transcript) she is rather using a question than a statement to indicate uncertainty and to emphasize her non-expert-identity. She states that *you drink a lot of beer* and is using a common stereotype for Germans. While avoiding stereotypes during the conversation so far, she finally gives in and uses a stereotype herself. The context suggests that she uses *you*, the second person plural pronoun, meaning the Germans as a whole and not *you* as the interlocutor. The use of the second person plural pronoun indicates that she includes her interlocutor into the German culture. She positions herself as a non-German who attributes the stereotype to the German culture and views her interlocutor as being part of that German beer drinking culture even though she cannot be sure if he is really a part of it or not. She uses laughter in combination with her suggestion of beer to create rapport and signal politeness and to some extent non-seriousness. She is joking about her bringing up a stereotype like that, which creates common ground and establishes a good relationship between the interlocutors as she is indicating by the laughter that she is not insulting her conversation partner by using a stereotype. The conversation goes on about different kinds of beer and about differences and similarities of beer in Germany and the U.S., which also underpins the good relationship and the equal level the participants talk about cultural differences.

Example 12 (fish and chips) is another example that is concerned with **stereotypes**, in this case with the British stereotypical food *fish and chips*. It shows very well how the participants talked about stereotypes and how they discussed them without judging and instead make fun and joking about the stereotype by using laughter to create rapport. It also shows that stereotypes do not necessarily have

to be bad, but that they could be used to create common ground in a very positive way if the discussion about them is accompanied by laughter and held in an intercultural sensitive way.

(30) Differences and similarities – Christmas (see example 14)

In example (14) (Christmas), the participants discuss **differences and similarities** with regard to the same concept, in this case Christmas. They talk about personal or family traditions and compare it to what the other person usually does and thus draw conclusions from their own behavior to cultural similarities and differences from their respective cultures. They come up with differences and similarities and accept these differences and similarities without judging, which is a very cultural sensitive way to talk about cultural differences. They both show openness and curiosity with regard to their interlocutor and the way they are celebrating Christmas.

The following example is concerned with showing sensitivity.

(31) Showing sensitivity – haunting (06SB157BO11; #06:30#)

- SB157: [<and> apart from that u:hm, #00:06:30#
for /Christmi/- **for Christmas we usually have deer**,
or uhm [turkey]. {raises eyebrows}
- BO11: o:h?
- SB157: ((hehe)).
- BO11: really.
do you get-
can you get deer like at stores in Germany?
- SB157: yeah >well< it's a more of a special u:hm meat of course.
like we wouldn't have that every day,
... obviously so just for Christmas or special events,
I'm not really fond of it to be honest,
I'm not- I'm [not that],
- BO11: [ah].
- SB157: fond of meat in the first place but u:h,
... it's okay w-
if my grandma prepares it it's-
it's bound to be good ((hehe)). #00:07:00#
- BO11: oh wow.
very cool.
do you-
I mean .. you can ... get deer in Idaho but you have to ... shoot it.
[like ((laughing)) you have to go hunting ((laughing))].
- SB157: [o:h yeah right].
very different (((chuckles)))].
- BO11: [()] ((laughs)),
but uh yeah,
yeah I mean you **have to have like a license** and everything to go,

and a tag to go and get a certain ... kind of deer,
of a year something,
and then people () deer.
but that's really the only way you can get it #00:07:30#
I mean you can't like buy it in any way.
(I never have) [.h].

SB157: [oh] do you have any hunters in your family ((thh)) as well.
BO11: I'm- ... not really in "my" family,
u:hm ... but yeah again **it's like very popular sport**,
like in I- so it's like a sport °kind of° thing [[[laughs]]].

SB157: [[[laughs]]].
BO11: like lots of people,
lots of people (hunt).
there is like bow hunting and [°all kinds°].

SB157: [u:h ... interesting].
BO11: all kinds of °like yeah°.
very outdoorsy state ((hehe)) ya. #00:08:00#

SB157: yeah apparently.

In example (31), a very **cultural sensitive** topic is being discussed: hunting. SB157 brings up the topic by mentioning that she and her family are usually eating deer for Christmas. BO11 is surprised that one can buy deer in Germany, using *oh* as an interjection to express surprise and astonishment and asking a follow-up question (*can you get deer like at stores in Germany*) to clarify meaning. She compares it to what it is like in Idaho, the state she lives in, and states that you cannot buy deer in a grocery store, *you have to shoot it* [the deer] *if you want to eat it*. She is very carefully describing that you can also not simply shoot deer, but that there are certain safety restrictions (*license, a tag to go shoot*). She also explains that hunting is a *very popular sport* in the U.S. and SB157 asks a follow up question, wanting to know whether BO11 has hunters in her family in order to indicate interest. Idaho and other more rural areas in the U.S. are known for hunting and SB157 probably also draws on socio-cultural orientation knowledge by asking the follow up question, assuming that BO11 has family members that go hunting. They talk very carefully about the topic and it is very clear that they show sensitivity not only towards cultural aspects – like the differences in hunting – but also towards the topic itself, which has to do killing animals. They use hedges and laughter to mitigate the sensitive topic and to create rapport when comparing the concepts, but they are very serious when talking about the topic itself, which shows that they can adapt to the seriousness of the topic, which also shows cultural sensitivity.

The following teaching tasks are made-up tasks that should help establishing intercultural competence in the EFL classroom.

Teaching Tasks

Task 1: Look at the examples (27-31). What happened? How did the participants feel? How would you react? What can you learn from this story? Discuss your findings with your partner.

Task 2: Imagine you are taking part in a school exchange. Partner A is the German exchange student and Partner B is the exchange student from the foreign country. Partner B and his/her family are serving you their favorite food – a typical dish from their region. Unfortunately, it is made from ingredients that you do not like. How would you react? Make up a possible conversation and act it out. (afterwards: feedback from classmates)

Task 1 is based on a task in the DESI-IKK assessment, which assesses the specific intercultural competence with the help of critical incidents (see section 2.1.5.2.1. and 5.2.). The questions that I chose for this task are similar to the ones in this DESI-IKK assessment (see Göbel 2011: 203; see also Hesse & Göbel 2007: 265), which is based on Bennett’s development model of intercultural sensitivity (see section 2.1.5.2.1. and 5.2.). The first question (What happened?) draws on a cognitive situation analysis, the second question (How did the participants feel?) is an affective situation analysis. Question 3 (How would you react?) is a question that deals with strategies to act and react and question 4 (What can you learn from this story?) is a transfer and evaluation task that the learners should complete in order to reflect on the situation and draw conclusions for application in own intercultural encounters. These questions address different cognitive levels and different levels of difficulty, which makes this task very diverse. Examples 27 until 31 can serve as “critical incidents” (see Lüsebrink 2016; Spencer-Oatey & Harsch 2016), but in a positive way and as positive and successful examples, not in the classic definition of intercultural failure. After the (inter)cultural analysis, learners discuss their findings with their partners so that they have to explain and argue what they came up with, in order to establish communicative competence as well.

Task 2 is a TBLL task, in which the learners should imagine they are taking part in a school exchange and thus simulate an intercultural encounter. This example is critical situation in which the German exchange student (partner A) does not like the food his or her American exchange student and his or her family are serving him. The learners should show how they could react in this situation in an interculturally sensitive way. They have to think creatively and make up a dialogue, while at the same time drawing on their problem-solving competence, social competence as well as intercultural competence. It should be noted that task 2 would also work with two different groups for partner A in order to show more variety. Partner A in group A could for example act in an ethnocentric way, and partner A in group 2 could act in an ethnorelative way. The classmates would have to comment on the behavior and come to the conclusion that the ethnorelative way is better in the sense that it is more culturally sensitive and competent.

Both tasks are high intermediate or advanced tasks as they are very complex and draw on higher levels in Bloom's taxonomy, namely level 5 (synthesis) and 6 (evaluation).

Another teaching task example is provided as a case study on how to explain a (new) concept (see section 5.5.2.). In the case study, I will provide a full lesson plan, lesson objectives and material in order to illustrate how "explaining a new concept" could be taught in the EFL classroom and thus how intercultural competence can be established. These examples and the case study could be used directly in the EFL classroom.

5.4. MULTIMODAL FEATURES OF TACoCASE

One of the key benefits of TaCoCASE is its multimodal setup. The conversations are video-mediated and thus can be classified as CMC conversations (see section 3.1.1.). The CEFR describes **online interaction** as one interaction activity that had been added in the latest CEFR version, which makes it a fairly new aspect to investigate. Online interaction is defined in the CEFR as communication

“mediated through a machine” (CEFR 2020: 84), which classifies it as CMC (see section 3.1.2.2.). Integrating online interaction into teaching is necessary, because it differs from face-to-face communication – in fact, the CEFR itself emphasizes that online interaction “is unlikely ever to be exactly the same as face-to-face interaction” (CEFR 2020: 84). Online interaction plays an important role in a globalized and digitized world and makes up a large part of communication in our everyday life. Learners need to know about new aspects and strategies that come with online interaction, which means that online interaction should be integrated into teaching.

Another aspect that can be analyzed due to the multimodal setup in TaCoCASE is **non-verbal communication**, as non-verbal elements are transcribed in detail (see section 3.1.2.3.2.). Non-verbal communication, however, is an aspect that is not covered in EFL teaching (see also Byram 2021: 18). According to Byram, “non-verbal communication is not included in a model for teaching in foreign language classroom, even though its significance is fully acknowledged” (Byram 2021: 36). This section tries to close this gap and provides examples of online interaction and non-verbal communication that could be used in the classroom.

In order to teach multimodal features and aspects of online interaction and non-verbal communication, I established the following objectives that should be reached by the learners.

The objectives for learning special features of TaCoCASE are:

1. The learners can find strategies to communicate successfully online and with non-verbal elements.
2. The learners can develop requirements for successful online communication and non-verbal communication.
3. The learners can apply and evaluate their strategies and requirements for successful online and non-verbal communication.

5.4.1. ONLINE INTERACTION

With the rise of globalization and digitization, especially in post-Covid-19 times, **online interaction** has gained much importance in our every-day life and in business and work settings. Online interaction can be characterized as computer-mediated communication (CMC) as it usually involves either a **computer** or at least **video-mediated communication** via a technical device. CMC and its characteristics have already been covered in section 3.1.2. in greater detail.

Online interaction in the CEFR

Online interaction is mentioned in the CEFR in the updated version (2020). It is centered around the question how interlocutors communicate online and the strategies or concepts that should be known to students are operationalized. Online interaction is mentioned in the CEFR as an important aspect of interaction. The CEFR lists two reasons for using online communication: serious issues and social exchanges (CEFR 2020: 84). These situations are real-life situations that the learners could encounter in their lives and thus, EFL teaching should prepare learners to be able to act competent in these situations. The following key concepts for online interaction were described by the CEFR:

Key concepts operationalized (CEFR 2020: 84):

- instances of simultaneous (real-time) and consecutive interaction, the latter allowing time to prepare a draft and/or consult aids;
- participation in sustained interaction with one or more interlocutors;
- composing posts and contributions for others to respond to;
- comments (for example, evaluative) on the posts, comments and contributions of others;
- reactions to embedded media;
- the ability to include symbols, images and other codes to make the message convey tone, stress and prosody, but also the affective/emotional side, irony, etc.

The list of key concepts suggests that the CEFR mainly focuses on written or asynchronous CMC, which is only one aspect of CMC (see section 3.1.2.2.4.). Synchronous CMC or real-time online interaction as in Skype conversations or other video-mediated conversations, is not specifically taken into account as far as the key concepts are concerned. Examples from TaCoCASE, which count into the category of

synchronous CMC, could serve as good examples to establish more key concepts. Knowing how to deal with a communication breakdown due to technical issues, how to deal with echo or interruptions , to name a few, would be additional factors that should be included. I would highly suggest to integrate these key concepts for synchronous online interaction in the CEFR as well and thus looking at the broad range of CMC and not only one aspect of it.

In this category of online interaction, not only key concepts are operationalized; the CEFR also lists factors and requirements for successful online communication (see figure 58). These factors count for both asynchronous and synchronous CMC, and can therefore be used as criteria for successful online interaction as discussed in this work with TaCoCASE and the Skype conversations.

Requirements for successful online communication (CEFR 2020: 84):

Successful online communication
• the need for more redundancy in messages; • the need to check that the message has been correctly understood; • ability to reformulate in order to help comprehension and deal with misunderstandings; • ability to handle emotional reactions.

FIGURE 58: SUCCESSFUL ONLINE COMMUNICATION (CEFR 2020: 84)

Some of the requirements are due to the setting of online interaction and the mediation through a computer or technical device. Understanding problems could arise through technical issues or the lack of a stable internet connection. Other factors, such as overlap for example, are due to the technical setting and sometimes also due to non-simultaneous comprehension. The technical setting, the internet connection and the distance between the interlocutors contributes to a slightly different interpersonal relationship between the interlocutors than in real-life and on-site interaction. Knowing these limitations, being aware of possible challenges and applying the strategies mentioned above,

make online communication successful. To what extent some of the factors contributed to successful communication in TaCoCASE, will be explored with the help of the following examples.

Selected examples from TaCoCASE

(32) Online interaction: Redundancy – reality TV shows (08BO15BI36; #15:01#)

BI36: [I find that reality TV shows are so different].
BO15: [(ehh) ()].
BI36: hm?
BO15: oh I'm sorry?
BI36: oh.
 (I find that) reality TV shows are so different, #00:15:01-4#
 .. here in comparison to .. America,
 like I've seen some,
 ... and I've seen some here,
 [and it's],
BO15: [oh it's],
BI36: i- ... it's different.

In example 32, the participants are talking about TV shows and the British speaker states that *I find that reality TV shows are so different*, when comparing reality TV shows in the UK and the United States. Due to overlap and apparently a technical issue, the American speaker does not understand what the interlocutor is saying and asks for clarification (*oh I'm sorry?*). The British speaker uses this question as a prompt to repeat the exact same phrase *I find that reality TV shows are so different*. She uses **redundancy** of the expression to make sure that her interlocutor understood what she said and thus in order to secure understanding.

(33) Online interaction: Need to check that the message has been correctly understood – Can you hear me? (04BO16BI40; #01:00#)

BO16: ... oh- okay.
 this is biza:rre.
 fai:led .. played test tone.
 this is .. °NOT good°.
 .. okay.
 give me just a moment please.
 I'm so:rry:.
BI40: that's alright.
BO16: (4.2) (°that sucks°.)
 AH.
 can you hear me?
BI40: .. yeah. {drinks from a bottle}

... I can hear you.
 BO16: [okay].
 BI40: [(ehh)] h. {recaps bottle}
 BO16: ... THAT was really bizarre.
 .. OKAY.
 .. HELLO:?
 BI40: . . hi?

In example 33, the American speaker has problems with his connection or with the program that he is using (*failed .. played test tone. this is .. °NOT good°.*), so he asks *can you hear me* several times to see whether the British speaker understood him. The context suggests that he feels the **need to check if the message had been understood correctly** due to his technical issues. The British speaker signals understanding by saying *yeah I can hear you* and they can finally hear each other, which the American speaker resolves by using *okay* as a backchannel and comprehension signal. He evaluates the technical issues and the non-understanding by saying *THAT was bizarre* before starting over the conversation and transitioning into a new topic. He uses *okay* as a closing marker of that topic section and as topic starter and transition marker to the new topic, namely the re-start of the conversation, which he introduces with a greeting (*hello*). This whole non-understanding phase is accompanied by many small pauses and even one four seconds pause, in which the American speaker is trying to get the technology to work. Due to the fact that the American speaker explains that he has issues (*failed played test tone; this is not good*) and asks for a moment to fix it (*give me just a moment please*), the American speaker builds rapport and provides information so that the British speaker knows what is going on in order to clarify the problem. The context suggests that these strategies are successful as the relationship between the speakers seems not to be affected by the technical issue. This is a nice example that shows how to successfully cope with technical issues and prevent a communication breakdown.

Teaching Tasks

Task 1: What are differences or similarities of online vs. face-to-face communication? What challenges and opportunities arise for online communication? Make a list and compare your results with your partner.

Task 2: Look at the examples (32-33). What challenges (or opportunities) do the participants face? Identify and explain the strategies they used in order to successfully communicate (online)? (Apply these strategies yourself in an online conversation with an exchange student.)

In **task 1**, the learners should identify differences and similarities between online and face-to-face communication before comparing the results with a partner. It is a task that draws on two cognitive domains in Bloom's learning objective taxonomy: In the first part – identifying differences and similarities of online and face-to face conversation – it draws on level 2 (comprehension) and in the second part – comparing the results – it draws on level 4 (analysis). Since some technical vocabulary is needed in order to compare online and face-to-face communication, I would propose to use this task for intermediate or advanced classes.

Task 2 uses the corpus examples from TaCoCASE and the learners should identify and explain (level 2 of Bloom's taxonomy) the strategies the speakers used. As an additional step, the learners could apply the strategies themselves in an online conversation in an online exchange (if the school or class takes part in exchanges like that). Both tasks make aware of strategies in online interaction and task 2 even encourages learners to take part in online interaction themselves. Including online interaction in EFL teaching is quite new approach and not very common, which makes these tasks a nice and forward-thinking contribution to EFL teaching in general.

5.4.2. NON-VERBAL COMMUNICATION

Non-verbal communication plays an important role in (online) interaction and in multimodal meaning making. Michael Argyle (1988), coming from a social psychologist perspective, summarizes the research on non-verbal communication in his field and comes up with several dimensions of non-verbal communication and the signals that represent non-verbal communication.

Signals of non-verbal communication (Argyle 1988: 1):

- Facial expression
- Gaze
- Gesture and other bodily movements
- Bodily posture
- Bodily contact
- Spatial behavior
- Clothes and appearance
- Non-verbal aspects of speech

These aspects, that Argyle mentioned, are very well established in SLA research and still up-to-date. They appear in spoken language and in interaction – online as well as offline. Non-verbal communication is an important element in the negotiation of meaning in a multimodal setting and contributes to mutual understanding. According to Argyle, non-verbal communication fulfills four main functions: expressing emotions, communicating interpersonal attitudes accompanying and supporting speech, self-representation and rituals (Argyle 1988: 5). In our transcription guidelines for example, we also include the signals mentioned by Argyle (see section 3.1.2.3.2.) and we find that gestures and facial expressions occur very frequently in our data. In a more holistic approach to non-verbal communication, Goodwin (2000) and Mondada (2014) for example look at embodiment and gestures in meaning making in conversation. Goodwin (2000) argues that language and the environment cannot be looked at in isolation of each other and suggests to look at “simultaneous use of semiotic resources by participants” (Goodwin 2000: 1490). Mondada (2014: 95) looked at pointing gestures in particular and views them as “multimodal resources deployed by the speaker”. This multimodal approach to analyzing language use seems to be very successful and this approach is also undertaken in this work. Gestures have also been subject to analysis in learner language research. Dimoski and colleagues (2016) found that learners use gestures as effective communication strategies (Dimoski et al. 2016: 75). Studies also found that learners used gestures more frequently in their L2 than in their L1, which underpins the argument of gestures being an effective communication strategy and a way to create meaning (see Gullberg 1998, Zhao 2006).

Despite these research findings and the importance of non-verbal communication in interaction, non-verbal communication is usually not subject to foreign language teaching. Foreign language teaching simply does not cover it and it is “not included in a model for teaching in foreign language classroom, even though its significance is fully acknowledged” (Byram 2021: 36). Poyatos (1992) also argues for the integration of non-verbal communication in the EFL classroom as speech does not only consist of language, but also paralinguistics and kinesics. There seems to be a gap between authentic spoken interaction, including non-verbal communication, and interaction or communicative competence that we teach in the EFL classroom (Byram 2021).

Non-verbal communication in the CEFR

Non-verbal communication is only briefly mentioned in the 2001 version in the CEFR and it is divided into practical actions and paralinguistics (see figure 59). Practical actions include pointing, demonstrations as well as clearly observable actions (CEFR 2001: 89). Paralinguistics are again divided into body language, such as gesture or facial expressions, extra-linguistic speech sounds and prosodic qualities, such as voice quality, pitch, loudness and length (CEFR 2001: 89).

Factors of non-verbal communication (CEFR 2001: 89)

Practical actions	Paralinguistics
• Pointing • Demonstration • Clearly observable actions	• Body language • Gesture • Facial expression • Extra-linguistic speech-sounds • Prosodic qualities • Voice quality • Pitch • Loudness • Length

FIGURE 59: NON-VERBAL COMMUNICATION IN THE CEFR (CEFR 2001: 89)

Due to the special transcription scheme (see section 3.1.2.3.2.), these features mentioned in the CEFR can easily be retrieved from TaCoCASE as all of them are transcribed in greater detail in TaCoCASE. The

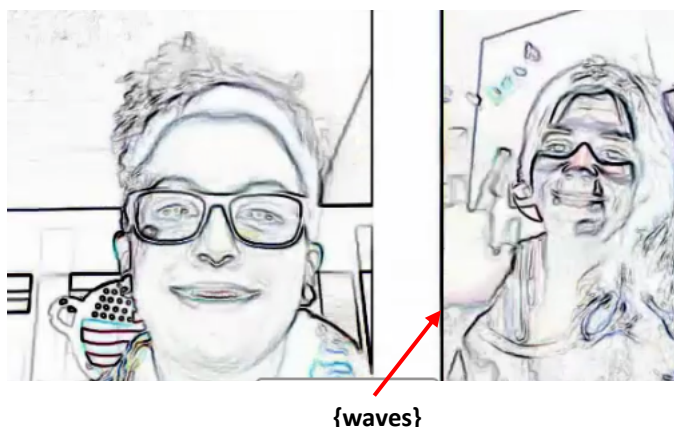
learners could watch the conversations or read the transcripts in order to find instances of non-verbal communication in the corpus. They could look at the instances and identify strategies the speakers used in TaCoCASE to combine verbal and non-verbal communication. In a further step, the learners could apply these strategies and use non-verbal communication themselves in interaction.

I will present some examples from TaCoCASE, in which the speakers used non-verbal communication in the following.

Selected examples from TaCoCASE

(34) Non-verbal communication: Gestures – waving (08SB147BI29; #40:00#)

SB147: "okay",
 BI29: yeah, {moves camera}
 SB147: "by:e"? {waves}
 BI29: yeah it was nice (°talking to°) you,
 {waves}
 SB147: .. yeah was nice talking to YOU,
 bye. {waves & smiles}
 BI29: >yeah<,
 "by:e". {waves & smiles}
 (2.4) #00:40:18-5#



In example (34), the speakers are at the end of the conversation in the closing scene. They say *bye* to each other and support this by using gestures and facial expressions to create rapport. In fact, the German speaker starts **waving** and the British speaker waves back and smiles at the German speaker ({waves & smiles}). The waving gestures and smiles accompany this whole closing scene as the speakers are waving and smiling and saying goodbye to each other all at once. The gestures support verbal communication and express emotions, which helps to create rapport leaves the participants with a positive feeling at the end of the conversation.

(35) Non-verbal communication: Pointing / showing – popcorn bag (03SB149BO03; #07:30#)

SB149: our main topic is food, is there anything that you like ...most? (1.5)

BI31: >I don't know?< ((clears throat))
I- I like.. {shows a bag of popcorn}
_popcorn.
popcorn (and me:).
[[((laughs))] {moves popcorn bag from side to side}

SB149: [[((laughs))]

BI31: Actually ((laughing)), actually I got my food for me
((laughing)). {moves hand to face}
>I like _popcorn<
u:hm

... I like these cakes. {shows a cake}
I <love> this ((/?i?z/)) cake (my mom) ().
(Because I) live at home, because I'm just a student
((ehh)) [.h] {moves hand to neck}

SB149: [yeah].

BI31: so.. I like.. {shows cake} uh these..()
cakes are so cute.
[[((laughs))] {holds hand in front of mouth}

SB149: [[((laughs))]



{shows a bag of popcorn}



{shows a cake}

In example (35), the participants are talking about their favorite food and as the British speaker has some of her favorite food with her, she is showing it to her interlocutor. The British speaker (BI31) says that she likes popcorn (*I like popcorn*) and before saying the word *popcorn*, she holds a bag of popcorn into the camera and **shows a bag of popcorn** ({shows a bag of popcorn}) to strengthen her utterance and signal importance. She is laughing and at the same time moving the bag of popcorn from side to side ({moves popcorn bag from side to side}). The context suggests that this could have two possible reasons: (1) she is either not sure whether her interlocutor can see the bag alright and is moving it in front of the camera from side to side to make sure that he can see it and to get the interlocutor's attention (see Mondada 2014) or (2) she is very proud and she finds it funny that she happens to have a bag of popcorn with her and therefore moves the bag from side to side to strengthen her utterance and signal importance. She goes on and says that she likes *these cakes* and refers to a cake that she is showing into the camera, introducing the referent (see Mondada 2014) ({shows a cake}). She explains that her mother makes these cakes at home, which is an act of affective stance-taking when providing

personal information and describing her emotional involvement. It is interesting to see that in the first instance, non-verbal communication is used to support the verbal utterance and to create rapport in a humorous way (laughter is involved) as the meaning would be clear just from hearing what the British speaker said and without showing the bag of popcorn. In the second instance, the showing is necessary as the verbal utterance only refers to *these cakes* and without showing the cake into the camera, the interlocutor would not know what kind of cakes the speaker had in mind. The German speaker does not say very much in this sequence, he is signaling understanding and prompting exploration through laughter non-verbally and verbally through using the word *yeah*, in order to mark comprehension.

(36) Non-verbal communication: Extra-linguistic sounds – syntax trees (02SB152BO06; #06:30#)

SB152: yeah we've got,
 we have to paint trees,
 with,
 BO06: oh ya.
 SB152: ...verb [and] whatever.
 BO06: [yeah].
 SB152: and [ya].
 BO06: [syntax] trees?
 ((laughs))
 SB152: yeah?
 ((thh))
 it's making me crazy actually. **((laughing))**
 BO06: mhm.
 SB152: nobody really explained how they
 work so,
 BO06: ahh.
 SB152: **((laughs))**
 we have to figure it out by ourselves,
 I think.
 [((laughs))]
 BO06: **[mhm]**.

Example (36) shows very clearly how **extra-linguistic sounds**, such as laughter are used. The German speaker is talking about her linguistics class and is complaining that they have to paint syntax trees. She uses an evaluative statement (*it's making me crazy actually*) to show her discontentedness. She strengthens this verbal utterance with non-verbal communication – a sound **((thh))** (clicking her tongue) – in order to underline and reinforce her statement. Her interlocutor uses another extra-

linguistic markers [*mhm*], to signal approval, confirm understanding and as the context suggests also to express empathy.

Teaching Tasks

I will present some possible teaching tasks in the following, in order to illustrate how non-verbal communication can be subject to teaching in the EFL classroom. As discussed earlier in this section, non-verbal elements are usually not taught in the EFL classroom even though their significance and importance in interaction are acknowledged (Byram 2021: 36). These tasks close this gap and provide authentic examples (from the corpus) that help making aware of the use and functions of non-verbal elements (NVEs).

Task 1: Watch the conversation (example 34-35) and write down all non-verbal elements the participants use.

Task 2: Choose two non-verbal elements from the previous conversation. Explain how they were used and why the speakers used them.

Task 3: Write down a conversation two people are having. Try to include at least five non-verbal elements. Make sure the NVEs are used in an authentic way.

Tasks 1 and 2 use TaCoCASE and examples from the corpus (example 34 and 35) to make aware of the use of non-verbal communication in combination with verbal communication. In **task 1**, the learners should find and write down the non-verbal elements they can find in the examples. This task has two challenges: Reading a transcript, which the learners are not used to and (2) identify the non-verbal elements, which are annotated in the corpus with curly brackets. Due to these challenges, I would recommend to use this task for intermediate or advanced learners, even though non-verbal elements in general should play a role at the beginner level too. At the beginner level, when the vocabulary range is not very broad, non-verbal communication can help the learners to express themselves when verbal communication cannot take place due to lack of vocabulary. As far as the cognitive domain is concerned, task 1 corresponds to level 1 (knowledge) in Bloom's taxonomy (see section 2.1.1.).

Task 2 is a follow-up task of task 1 and corresponds to Bloom's second level (comprehension). The learners should choose two examples of non-verbal elements and they should explain the use of these NVE to their partner. Task 2 is designed as partner work and as a communicative task, in which the learners have to explain the function of the NVEs. Task 3 is a written task for learners to complete individually. The learners should make up and write down a conversation, in which they make use of non-verbal elements. This task is a creative task that corresponds to level 5 (synthesis) in Bloom's taxonomy of cognitive involvement and educational objectives.

5.5. CASE STUDIES: HOW TaCoCASE CAN BE USED IN THE EFL CLASSROOM

The previous sections looked at aspects of intercultural competence as demanded by the CEFR and matched these aspects with examples from the intercultural encounters in TaCoCASE. In this section, the application of TaCoCASE in the EFL classroom is even taken further, which means that I established **two case studies** to show in an even more concrete way how TaCoCASE can be used in the EFL classroom. I picked two case studies, following the aim and rationale of this work to establish communicative and intercultural competences. The first case study (section 5.2.1.) focuses on **teaching vocabulary with TaCoCASE** and thus draws on linguistic competence in particular and communicative competence in general. The second case study presented in this work (section 5.5.2.) focuses on establishing intercultural competence by looking at how learners can **explain a (new) cultural concept**. These case studies demonstrate how TaCoCASE can be used to establish communicative and intercultural competence, alongside with digital competence as elements of data-driven learning are included as well.

The case studies follow a learner-centered and action-oriented approach and make use of cooperative learning (CEFR 2020; Decke-Cornill et al. 2015; Bonnet et al. 2017). In both case studies, an inductive and bottom-up approach is used, which means that the learners first look at authentic examples from

the corpus and then draw conclusions to come up with rules or strategies that help the learners establishing communicative and intercultural competence (Decke-Cornill et al. 2015). The first case study uses DDL and is a very hands-on corpus activity in order to acquire new words and thus establishing linguistic competence. The second case study is a TBLL activity that draws on mediation strategies and establishing intercultural competence. The case studies and the material are based upon and make use of the findings in chapter 4 and the previous sections of this chapter. Learning objectives as well as a lesson plan and teaching material (worksheets) will be presented for each case study in order to be very precise and authentic in the application. It is important to note that the lesson plans and the material depict only one possible teaching methodology and they could be adapted at any point.

In sum, the case studies reflect a very modern approach of establishing communicative, intercultural and digital competences as the methodology is based on state-of-the art teaching methods and the examples are carefully selected based on the evidence-based findings in this work.

5.5.1. HOW TO TEACH VOCABULARY WITH TACOCASE

The following lesson is designed as a DDL activity with hands-on use of corpora to create a vocabulary profile and thus **establish linguistic** as part of **communicative competence**. Additional worksheets help the learners to complete the DDL activity. The creation of a vocabulary profile is drawn on the concept of lexico-grammatical profiles of words (see section 5.1.4.; see also O’Keeffe et al. 2007: 14). It could as well be carried out as hands-off activity (without a computer) which means that the frequency list, collocate lists and respective concordances would have to be provided as screenshots on worksheets. However, the preferred methodology would be to use the computer and the corpora in order to establish corpus literacy and digital competences (see section 2.1.6.3.). The main competence that is drawn upon in this case study is communicative competence and in particular linguistic competence (vocabulary range) (see also section 5.1.1.).

Case study 1: Vocabulary profile

Grade: 8-9 (depending on corpus literacy)

Duration: 90 minutes (2 lessons)

Competences: Communicative competence (linguistic competence) and digital competence

Skills: Communication, collaboration, (creativity, critical thinking)

Main objective: Learners acquire a new word and create a vocab profile.

Specific objectives:

Learners can ...

- ... understand and explain the meaning of a word.
- ... find collocates and lexical chunks of a word.
- ... know about syntactic and semantic restrictions of a word.
- ... use the word in a sentence and in different contexts and situations.

Lesson plan:

Time	Phase	Task / Content	Interaction form / method	Material
5'	Intro	Different sentences with synonyms of <i>know</i> are presented. - What do the synonyms have in common? Find a synonym that could replace the verbs → <i>know</i> Topic: Vocab profile of <i>know</i> . - Aim: find out: What does <i>know</i> mean? (Different meanings?) How is <i>know</i> used in a sentence?	Teacher presents sentences on a screen / board (synonyms are underlined). DDL	Screen / board
45'	Development	Learners create a vocab profile of <i>know</i> and follow the instructions on the worksheet (Vocab profile task). - Read the instructions on the worksheet - Apply the corpus searches (steps 1-4) - Write down findings / hypothesis	DDL using WebCorpLSE Partner work	PC / Laptop WebCorp LSE Worksheet (figure 60) Worksheet (figure 61)
30'	Application	Learners discuss the findings / hypothesis with another group in exchanging the vocab profiles. Vocab profiles will be discussed in class.	Discussion I (group work) Discussion II (class)	Worksheets

10'	Conclusion	Learners talk to each other using different expressions including <i>know</i> or different verb forms of <i>know</i> .	Double circle	none
-----	-------------------	--	---------------	------

The **introduction phase** starts with a presentation of sentences that contain synonyms of *know*. The learners find out that the correct synonym is *know*. In that case, the learners already know synonyms of the word and can predict the meaning of *know*. The teacher then presents the topic of the lesson: Creating a vocab profile of the word *know*. The learners have already worked with corpora and the concept as well as elements of a vocabulary profile are clear to them. In the **development phase**, the learners follow the instructions on the worksheet (figure 60), which means they first read the instructions, conduct the corpus searches and then write down their findings. During the DDL activity, they become researchers on their own, discuss hypothesis with their partners and conduct the corpus searches together with a partner so that they also practice social skills and communication as well as collaboration skills. Creating a vocab profile has many advantages: (1) The learners know the meaning of a word and how it is used in context. (2) They extract pre-fabricated chunks that they can easily integrate into their language use. (3) The concordances provide examples, which the learners could write down to remember how the word is used in context. The learners acquire the word in many different ways and thus save it in their mental lexicon with many connections to other words. The learners write down all the findings that they came up with on their worksheets (see figure 60). In the **application phase**, they compare the vocab profile with a different group and thus maybe add some more information to their vocabulary profile. Again, social skills as well as communicative and collaboration skills are trained. In the **conclusion phase**, the learners apply the knowledge and use the new word they have acquired. Using the double circle method (see Haß 2009: 12), the learners stand in two circles facing each other. They start talking to the partner they are facing and after one minute, the outer circle spins and the learners turn to another partner. The learners talk to at least three different people and thus need to adapt to the conversations and may think of even more examples. They also hear different examples as the interlocutors have different sentences and use the word

differently. The fact that the learners formulate sentences (production) and listen to their interlocutor (reception) draws on two language activities at the same time. The learners also establish interaction skills as they need to adapt to different partners in different conversations.

Please note: The activity would of course work with any other word. The learners could for example also make up a word they find difficult or they do not know.

Vocab profile



Go to WebCorpLSE and search for the word know. Create a vocab profile for know.



(<https://www.webcorp.org.uk/wcx/lse>)

To create a vocab profile, follow these steps:

1

Search for the word know in TaCoCASE. What POS is know? Is it a frequent word?

2

Collocates: What are words that are frequently used in combination with know (Collocates)? Find examples.

3

N-grams: What are chunks or idioms that contain the word know? Find examples.

4

Syntactic restrictions / semantic relations: How is know used in context? Are there any syntactic restrictions? Are there any semantic relations or connotations? Find examples.



Write down your findings and fill out the worksheet (vocab profile). Discuss your findings with your partner.



FIGURE 60: WORKSHEET VOCAB PROFILE - TASK

The worksheet is titled 'Vocab profile' in a large, bold, black font, centered at the top. Below the title is a yellow rectangular box labeled 'WORD'. Underneath this, there are several colored boxes for notes: a purple box labeled 'verb forms' followed by a light pink rectangular box; a purple box labeled 'meaning(s)' followed by a light pink rectangular box; a purple box labeled 'register' followed by a yellow rectangular box; a purple box labeled 'collocates' followed by a large teal rectangular box; and a purple box labeled 'chunks / idioms' followed by a light pink rectangular box. The entire worksheet is set against a light beige background.

Vocab profile

WORD

verb forms

meaning(s)

register

collocates

chunks / idioms

FIGURE 61: WORKSHEET VOCAB PROFILE

Vocab profile

(to) know (VERB)

verb forms

knowing (present participle), knows (3rd person singular present tense), knew (past tense)

meaning(s)

understanding, having information, being sure about something

register

formal & informal (depends on context)

collocates

- **even:**
Example: I don't even know → usually negative meaning
- **really:**
Example: I don't really know, there is no way of really knowing something → usually negative meaning

chunks / idioms

- you know (discourse marker)
- you know what I mean (discourse marker)
- I don't know
- I don't know what / if / how
- Do you know...? (Question)
- you never know

FIGURE 62: WORKSHEET VOCAB PROFILE (TEACHER VERSION)

5.5.2. HOW TO EXPLAIN (NEW) CULTURAL CONCEPTS WITH TACoCASE

This lesson is designed to **establish intercultural competence** and uses a TBLL activity and cooperative learning so that the learners have a goal-oriented and meaningful task – a simulation of a real-life situation – and an authentic problem they need to solve. The learners learn how to explain a (new) cultural concept, which means that they build up mediation strategies. Thus, the learners establish communicative and intercultural competence at the same time.

Case study 2: Explaining new (cultural) concepts

Grade: 9-10

Duration: 180 minutes (two double lessons)

Competences: Intercultural competence, communicative competence (discourse competence) and digital competence

Skills: Communication, collaboration, creativity, critical thinking

Main objective: Learners can explain a (new) cultural concept.

Specific objectives:

Learners can ...

... name discourse elements (discourse structure) for explaining a (new) cultural concept.

... extract and make use of expressions for each of the discourse elements when explaining a (new) cultural concept

... apply the discourse structure to explain a (new) cultural concept to an interlocutor.

Lesson Plan

Time	Phase	Task / Content	Interaction form / method	Material
Lesson I+II				
5'	Intro	Brainstorming: Traditional German food. Name examples of typical German food.	Class	Post-its / cards on board

		Imagine you are in XY and you talk about your favorite food? Do you think your interlocutors know the concept? What are possible challenges? Topic: How do we explain a (new) cultural concept?		
40'	Development	Objective: Find strategies to explain a new cultural concept - Listen to the example and read along. - Answer the questions: - What is the (new) cultural concept the speakers talk about? - How do the speakers talk about it? (discourse structure) - What strategies does the German speaker use?	Individual work	Worksheet (figure 63) TaCoCASE conversation (video)
45'	Conclusion	Compare the results with a partner. Fill out the worksheet and identify the discourse structures. Give examples for each of the discourse elements.	Partner work	Worksheet (figure 64)
Lesson III + IV				
5'	Intro	Recap of strategies to explain a new concept	Class	Worksheet (figure 65/64)
40'	Application	Objective: Apply the strategies to explain a (new) cultural concept. Learners pick a card with a food item. Group A: Käsespätzle / Group B: Schnitzel Task: You have an exchange student from the U.S. and you take him to a restaurant. He/she does not understand the word on the menu (A: Käsespätzle / B: Schnitzel). - How could you explain the food item to him/her? (research / strategies) - Imagine a conversation and act it out.	TBLL: Group work	Cards with word / picture of food item.
45'	Conclusion	Some groups will act out their conversation. Reflection / Feedback: What was difficult? What was easy? New strategies?	Group work / simulation Reflection in class / Feedback	none

The lesson plan is divided into **two parts**: Lesson I+II and lesson III+IV. In the first part, the learners are made aware of strategies how to explain a concept and they extract the strategies from the corpus example. In the second part, the learners themselves explain a new concept, applying the strategies they discovered before.

Lesson I+II: In the **introduction phase**, the teacher and the learners start out with a brainstorming activity about German traditional food. The learners name food and dishes that come to their mind and the teacher collects the terms on the board. Alternatively, the learners could write a food or dish on a card and the teacher could collect the cards on a flipchart or the board. The teacher asks the learners if people from a different culture would understand the items. The learners are made aware that they cannot take for granted that their favorite food is known in other countries and the learners anticipate possible challenges when explaining their favorite food. The topic of the lesson is introduced: How to explain a (new) cultural concept. The objective in the **development phase** is to extract discourse strategies or elements for explaining a new concept. The teacher prepared a worksheet (figure 63) with an example from TaCoCASE. The learners work individually and follow the instructions on the worksheet, which means they watch the conversation and read along. At this point, the learners use and build on their communicative competence in a receptive activity (watching the video / listening to what the speakers say). The learners answer the questions on the worksheet, which means that they identify the discourse elements of the German speaker's explanation of *Sauerkraut*. In the **conclusion phase**, the learners compare and discuss the strategies they identified with a partner. Working in pairs and discussing the findings draws on social skills as well as communicative and collaborative skills. The teachers prepared another worksheet (figure 65 or figure 66) that the learners fill in the with the identified and consolidated findings – strategies and examples. This worksheet serves as a strategy worksheet for the learners that they can use over and over again during their school career (and in real life). The strategies for explaining a new concept are clear at the end of the lesson. A possible solution of the worksheet is depicted in figure 64.

Lesson III + IV: In the **introduction phase**, the learners look at their strategy worksheet again and repeat the strategies for how to explain a (new) cultural concept. The focus in this lesson is on the application of these strategies. The learners work in groups in a TBLL activity. The groups are assigned different food items – Group A: Käsespätzle and Group B: Schnitzel. The teacher presents the task to the learners:

Task: You and your exchange student from the U.S. go to a restaurant. He/she does not understand the word on the menu (A: Käsespätzle / B: Schnitzel). How could you explain the food item to him/her? Imagine a conversation between the two of you and act it out.

This task is very open and the groups have to find a way on their own to solve the problem, which draws on their creativity and problem-solving competence. First, the learners need to come up with a strategy how to explain the concept. They could look at the strategies and do some research on the food item (how it is produced etc.). After they established a strategy, they need to imagine a conversation, which means that they have to do a perspective shift. On the one hand, there is the German student who explains the concept. On the other hand, there is the exchange student who does not know the concept. As far as the exchange student is concerned, the learners have to anticipate possible challenges and apply other mediation and interaction strategies in order to come up with an authentic conversation (for example asking for clarification, co-operating). The learners then should practice acting out the little conversation they made up. In the **conclusion phase**, the learners present their conversations in a simulation and the two concepts – Käsespätzle and Schnitzel – will be explained. The learners give peer feedback with regard to intercultural behavior, which means they look at cognitive, affective and action-oriented aspects and evaluate to what extent the explanation was comprehensible. The group itself could also reflect on the activity and explain what the challenges were. The learners thus apply strategies to mediate a new cultural concept and create common ground. They act as an intercultural speaker and establish intercultural communicative competence.

In a follow-up lesson, the cultural concept could be varied. The learners already practiced how to explain a food item. In a next step, they could practice explaining other cultural concepts in the field of customs and traditions or TV series ("Let's talk about..."). The worksheets (figure 66 and 67) could be used as a basis and the learners could again be in a double circle which facilitates communication and draws on all language activities at the same time (reception, production, interaction and mediation).

Please note that the worksheets are structured and designed based on the findings in the analysis.

How to explain a new concept

Listen to the example and read along.



TaCoCASE - 07SB149BI31

SB149: some- some typical,

like stereotypical German [uh:m]

BI31: [°yeah°]

SB149: things are good as well. (1.3)

so I don't know if you- if you've ever eaten ☐ Sauerkraut ((German (0.8)))?

or Kraut ((German (0.5)))? ...

BI31: no, what's that?

SB149: it's like <cabbage>? {lifts head up} .. A:[nd]

BI31: [O:H.]

SB149: that [is-] {moves hands in circle}

BI31: [°>okay.<°]

SB149: made (up) with some vinegar and stuff like that,

and it's uh: left inside a big box o:r a- a- a round (thingy)?

BI31: yeah, [a (thingy)]. {touches shoulder & pulls sleeve}

SB149: [y:nd] where ((laughing))-

where it is- where it stays,

and uh:m slowly {moves hands in circle}

.. turns into Sauerkraut ((German(0.9))),

so: the cabbage is.. slowly turning into that.

.h a[nd u:h.]

BI31: [>Sauer<]. ((German (0.3))) {leans forward & looks to side}

SB149: >yeah<.

.. h ..it's-

BI31: ((chuckles))

SB149: ((hehe)) it's like the- the typical stereo.. ☐ {looks to side & raises eyebrows}

typic thing that Americans say about Germans,

is that they are eating Kraut ((German (0.4))).

(1.5) so they also refer to us as the Krauts ((German (0.5))). ...

BI31: o:h it's {leans forward} [OKAY THAT'S [INTERESTING.] ☐ {looks down}



- What is the **new (cultural) concept** the speakers talk about?
- **How** do the speakers talk about it?
- What **strategies** does the German speaker use to explain the (new) concept?

FIGURE 63: WORKSHEET I - EXPLAINING A NEW CONCEPT - TASK

How to explain a new concept

Introduction of new concept / topic

- Introduce a new concept or a new (cultural) topic.
Example:
Do you know XY? / Have you heard of XY?

Check level of knowledge

- You need to find out whether you and your conversation partner have the same level of knowledge.
Example:
Yes, I've heard of it. / No, I don't know what it is. / What is XY? / I don't know if you know that. / I don't know if you've tried XY.

Evaluation*

*optional step

- You express your reaction to (not) knowing the concept.
Example:
Oh you should try it. / It's so good. / It's really cool / Oh, how do I explain it?

Explanation / Elaboration

- You share your experience, relate to known concepts and compare differences and similarities between your respective concepts. You explain the concept to your conversation partner.

Example:

- Spelling (B-R-E-Z-E-L),
- Describing how it is produced / made (it's made with vinegar)
- Breaking down the process/concept in many steps / pieces of information
- Linking to prior knowledge (it tastes like cabbage)
- Comparing to known concepts (it looks similar to XY)
- Using non-verbal elements (gestures)
- Using multimodal features (e.g. sending a picture in the chat)

Result

- You and your conversation partner come to an agreement. Your conversation partner, who learned about the concept, could express (dis)like or signal understanding. You could encourage your conversation partner to try it out (recommendation).

Example:

Oh, that's interesting. / Oh, I might try it. / You should try/watch it. / That's similar to XY then. / I think it's the same, just XY.

FIGURE 64: WORKSHEET II - EXPLAINING A NEW CONCEPT – STRATEGIES

How to explain a new concept

Introduction of new concept / topic

●
.....

Example:

.....

Check level of knowledge

●
.....
.....

Example:

.....

Evaluation*

*optional step

●
.....
.....

Example:

.....

Explanation / Elaboration

●
.....
.....

Example:

.....
.....
.....

Result

●
.....
.....

Example:

.....

FIGURE 65: WORKSHEET III - EXPLAINING A NEW CONCEPT (STUDENT VERSION)

Let's talk about...

...customs and traditions

Fesitvity XY

History

Example:

Food

Example:

Actions & practices

Example:

Date / Information

Example:

FIGURE 66: WORKSHEET LET'S TALK ABOUT... CUSTOMS AND TRADITIONS

Let's talk about...

...(TV) series & movies

Name of TV series or movies

Story

Genre

Medium

Reasons for watching it

Characters

Anything else?

FIGURE 67: WORKSHEET LET'S TALK ABOUT ... TV SERIES

5.6. BENEFITS OF TACoCASE IN A TEACHING CONTEXT

The following table displays the **benefits of TaCoCASE** in a teaching context at a glance. Table 45 is similar to the one in section 3.1.3., in which the general benefits of TaCoCASE were listed. The first and second column are the same and the third column describes the benefits of TaCoCASE with special regard to EFL teaching and its use in the classroom.

TABLE 45: BENEFITS OF TACoCASE IN A TEACHING CONTEXT

Key element	Description	Benefits for teaching / applications
Spoken language corpus	Spoken language features (e.g. overlap, intonation, hedges)	Spoken language; audio-visual input; reception activities (listening and audio-visual comprehension); communicative competence
CMC communication	CMC features (e.g. interruptions, echo), technical issues	Raising awareness of CMC features, strategies how to communicate successfully in online interaction (how to deal with interruptions or technical issues, building relationships in CMC communication etc.)
Multimodal set-up	Multimodal input: Audio and video: intonation, non-verbal elements (e.g. gestures, gaze, embodiment)	Communicative language competence (lexical, phonetic competence), non-verbal elements, multimodal meaning making; online interaction
Transcription	Different layers: orthographic, XML, etc. special guidelines: pragmatic elements (e.g. laughter), CMC elements (e.g. echo), ELF elements (e.g. code-switches), non-verbal elements (e.g. gestures)	Learning how to read transcriptions; analyzing different elements (pragmatic elements, non-verbal elements); reception: reading along the transcripts when listening to conversations (reading comprehension; listening comprehension)
Authentic language use	Video recordings of conversations; no lab-setting	Communicative language competence; authentic examples and not textbook examples; motivation as the topics are close to the learners' real world experience; participants are not much older than the learners, which means that their language use is similar
Informal register	Casual conversation via Skype between students; informal setting; prepared topic prompts but any topics can be discussed;	
Different speakers (native and non-native)	Native speakers and non-native speakers could be separated; TaCoCASE as reference corpus for ViMELF; ELF conversations (ViMELF plus TaCoCASE SB-BI and SB-BO)	Communicative competence; different varieties of English; "mistakes"
Intercultural encounters / intercultural communication	First encounters between students with different language and cultural backgrounds	Intercultural (communicative) competence
TaCoCASE + ViMELF	Combination of the two corpora -> larger database	Comparisons across the two corpora (e.g. language, discourse marker use)
Access	Video and transcripts available online (for free); integration in WebCorpLSE (searching the	Reception activities (reading and listening); Learner-centered approach (DDL) and teacher-learner interaction possible (depending on technical equipment)

	corpus online or comparing it with other corpora)	
--	---	--

TaCoCASE is a **spoken language corpus** that consists of conversations between students with different language and cultural backgrounds. Spoken language examples and spoken language features could be discussed in the classroom as audio-visual input from the corpus or in listening or audio-visual comprehension activities. Since the conversations were held via Skype, the corpus represents **computer-mediated communication**, which is an important aspect of communication that is not included in curricula or in the EFL classroom to a large extent even though CMC takes up an important part of communication in the learners' life. Knowing how to deal with CMC features like for example echo, or interruptions of recordings due to technical issues is important for the learners and they need strategies to communicate successfully online and to build relationships in CMC environments.

The data in TaCoCASE is audio, video and text (transcripts), which means that due to the **multimodal set-up**, EFL teachers and learners can look at either only visual elements like gestures, or a combination of modes like visual elements and language combined. This multimodal approach provides many benefits for learners and they could establish several competences at once and they learn that different modes can be used in combination, which contributes to multimodal meaning making. The **transcripts** of TaCoCASE could be analyzed by the learners as written input and they could be used in receptive activities (reading). Reading transcripts is a competence that the learners should also establish, especially in a more and more digitized world and as part of a corpus literacy. Learners could look at different layers of the transcript and look at the interplay of for example lexis, pragmatic elements and non-verbal elements.

Probably the most important benefit of TaCoCASE is that the conversations are real world language examples, which means that the corpus represents **authentic language** and examples from the corpus are authentic in contrast to many scripted dialogues in course textbooks in the EFL world. In terms of register, the conversations could be classified as informal as they are casual conversations about every

day topics. The examples from TaCoCASE are probably very motivating for the learners as the examples represent situations that are **close to the learners' real world experience** and the learners could also end up in situations like this, which makes the examples and the situations very relatable and authentic for the learners. The **informal register** is usually also not very much emphasized in the EFL classroom, even though informal conversation makes up a very large part of conversation in general, be it in conversation with friends and family or in small talk conversations in a business environment. Thus, it is important for learners to know about informal language and how to use it as well as choosing the appropriate register in a conversation.

The speakers in TaCoCASE are native as well as non-native English speakers, which means that the learners will be confronted with different **varieties of English**, which they are probably not used to. ELF conversations are usually not used in the EFL classroom and learners will establish a different take on different varieties of English and on 'mistakes'. Learners will experience that communication can be successful even though speakers make 'mistakes' and maybe pronounce words wrong or use wrong words. This experience could have an impact on the learners' communicative behavior as they could speak more spontaneously and fluently without thinking too much about possible errors which would normally hinder them to produce output.

Besides these communicative aspects, one of the main benefits of TaCoCASE is that the conversations are intercultural encounters and that examples from the corpus or excerpts from the conversations in TaCoCASE can be used to **establishing intercultural communicative competence** (see sections 5.2.2., 5.3. and 5.5.2.). If more examples are needed or if only ELF conversations should be taken into account, TaCoCASE can be combined with ViMELF as both were compiled under similar conditions and the transcripts and videos look very similar. Comparisons across these two corpora are possible, which could be interesting to look at for the learners.

TaCoCASE is **easily accessible**, which makes it easy to select examples and integrate them in EFL teaching. The videos and transcripts are available online and TaCoCASE is integrated in the corpus

search engine WebCorpLSE, which means that the learners could do DDL activities online and would not have to download software before using the corpus.

In sum, it can be stated that TaCoCASE is a **rich source for authentic teaching material** that can be used in the EFL classroom in various ways. The examples discussed above illustrate that TaCoCASE **fulfills the CEFR standards** with regard to **communicative and intercultural competence**. The corpus is very well suited for improving linguistic competence and pragmatic competence as well as teaching interaction and mediation strategies. In light of the lexical approach, using TaCoCASE for teaching vocabulary offers a new way of teaching vocabulary, that not only draws on communicative competence, but also on **digital competences** and **future skills** in general. In terms of explaining a (new) cultural concept or creating common ground, the selected examples can be used in the EFL classroom to establish intercultural competence. The **authentic examples** lead to a better understanding of the respective situation and the learners are made aware of intercultural behavior and can evaluate intercultural sensitivity. Due to its **multimodal set-up**, TaCoCASE can also be used to teach aspects of online interaction and non-verbal communication, which are two aspects that are usually not subject to teaching, even though they play an important role in multimodal meaning-making. Many examples focused on food, customs and traditions as well as popular culture and TV series and movies in particular. These are aspects of everyday life that the learners can relate to, which sparks their interest and leads to a **higher motivation** and investment in the EFL classroom.

In general, the inductive approach that is suggested in the teaching tasks when using corpus examples from TaCoCASE, depicts a learner-centered and action-oriented approach that is not only demanded by the CEFR and the German education standards, but that it is state-of the art and modern EFL teaching. The **case studies** show that the learners acquire communicative and intercultural competences and are involved in all language activities – reception, production, interaction and mediation – sometimes even all four are drawn upon at the same time. DDL activities and TBLL is

facilitated and the advantages that the use of corpora in the EFL classroom bring, can be traced from the examples discussed above. The most important aspect, however, is authenticity. The corpus examples expose the learners to authentic language in use and authentic intercultural encounters, which is very good input to learn from and which in turn contributes to authentic language acquisition and authentic language use by the learners.

6. CONCLUSION

Establishing communicative, intercultural and digital competences in the EFL classroom plays a central role to prepare learners to be a social agent in today's international, globalized and digital world. The world around us changes and we face many challenges like globalization, migration and digitization. EFL teaching needs to adapt to these challenges in order to prepare learners in the best possible way for their future life. The overall question in this dissertation was how EFL teaching can be modernized and upgraded in order to keep up with these challenges. I suggested corpora and data-driven learning as powerful tools to modernize EFL teaching. I investigated to what extent the use of corpora – and TaCoCASE in particular – is beneficial for EFL teaching and demonstrated how examples from the corpus can be used in the classroom. In fact, a corpus – TaCoCASE – was compiled that had been especially designed for the purpose of this work and the application in the EFL context. Linguistic and (inter)cultural analyses looked at the language in use in the corpus and how the participant negotiate meaning and mediate cultural concepts in authentic intercultural encounters. These findings were used to create an evidence-based selection of words and phrases as well as discourse and mediation strategies that should be taught in the EFL classroom. These examples can be used directly in the EFL classroom and they have proven to contribute to establishing communicative and intercultural competence through their syllabus fit. The two case studies illustrated how DDL activities can be used to establish communicative, intercultural and digital competences – all at once, which is very rare in EFL teaching. I provided new material, suggested and applied new methods and promoted a new mindset in order to upgrade EFL teaching.

As chapter 3 illustrated, TaCoCASE is a **new spoken CMC corpus** that allows research in various fields and that has many benefits that facilitate research in different areas. It can be used in the EFL classroom and it is a rich source for teaching material. As a spoken CMC corpus depicting authentic intercultural encounters between students with different language and cultural backgrounds, TaCoCASE can be used for analysis of language and culture. Even though the analysis of language and

culture is separated throughout this work, so that a detailed analysis of each is possible, language and culture are intertwined and should be thought together and taught in combination (Kramsch 1991, 1998a). TaCoCASE depicts authentic language in use and authentic intercultural communication and thus negotiation of meaning and concepts. Authenticity is one key aspect for integrating corpora in language teaching and this counts for TaCoCASE as well (Gilquin & Granger 2010). Typically, text books and examples in the EFL classroom are not very authentic, which could be made up for when using TaCoCASE (or other corpora). The topics in TaCoCASE are everyday topics which are close to the learners' world of experience and the participants are only a little bit older than the learner. The learners can therefore relate to the topics and the participants' situation, which in turn leads to a higher motivation and investment (Deci & Ryan 1993; Darvin & Norton 2015).

TaCoCASE depicts **actual language in use** and thus, aspects of communicative competence can be established with the help of authentic language examples. Using corpora for analyzing language and applying the findings for teaching communicative competence is quite common in applied linguistics research, where its full potential is acknowledged, but corpora are still not widely used in the EFL classroom (in Germany) (see Römer 2011; Mukherjee 2004; Jablonkai & Csomay et al. 2023; Pérez-Paredes 2019). However, using corpora for teaching aspects of intercultural communication, is a new approach that has not yet attracted much attention.

TaCoCASE depicts **CMC data and multimodal features** such as non-verbal communication and online interaction. How to communicate non-verbally and what challenges online interaction can have, is usually not taught even though it is crucial in multimodal meaning-making processes and in online collaborations in the learners' private or work life (Byram 2021).

The analysis of TaCoCASE (chapter 4) unveiled valuable insights in terms of linguistic and (inter)cultural aspects and thus in authentic language use and authentic intercultural encounters. As far as **linguistic aspects** are concerned, the language in the corpus was analyzed with corpus methods and the usual corpus searches have been conducted, such as frequency analysis, collocation and n-gram analysis as

well as concordance analysis (McEnery & Hardie 2012; O’Keeffe et al. 2007; Hunston 2002). The most frequent words in the corpus had been detected as well as frequent collocations and lexical bundles, which are useful to know and teach (Biber et al. 2004; Hutz 2018). Vocab profiles of the participants were created in order to find differences among NS and NNS in the corpus. No significant differences in the vocab profiles between the two speaker groups were found, which supports the hypothesis that the speakers in TaCoCASE can be regarded as speakers in an ELF intercultural environment and that there are only slight differences between NS and NNS as far as language proficiency is concerned. This also supports the idea that the (new) model in EFL teaching is not the native speaker any more, but the intercultural speaker.

Discourse markers, as pragmatic elements of language, had been analyzed as they usually produce difficulties for learners in terms of use (over and misuse). The focus in this work was on the DM *so* and *okay*, which are DM that appeared frequently in TaCoCASE and also appeared frequently together as collocates. Their use and functions were analyzed and it was found that *so* fulfills the following functions, which were the same basic functions as established in research: result and inference, action prompt, topic developer and sequencer, connection marker, marker of emergence from incipency and hesitation marker and floor-holding device. The analysis of the DM combination of *okay* and *so* revealed a sequence preference for *okay so* (instead of *so okay*) that appeared also across speaker turns. The function of the DM combination of *okay so* results mostly from the combined functions of the individual DM, but instances for new functions and function from one of the components could as well be discovered.

As far as **cultural aspects** are concerned, the analysis (see section 4.2.) showed that most of the conversations – in fact 2/3 of the conversations in TaCoCASE – are concerned with cultural aspects and that cultural topics play an important role in intercultural encounters. The cultural aspects under investigation were food, customs and traditions and popular culture such as TV series and movies. A semantic analysis identified words that are semantically related to food and that describe what, where and how food is made or eaten (see section 4.2.1.1.). The conversations about customs and traditions

and TV series were also investigated with regard to aspects the participants covered when talking about cultural issues. A corpus-based discourse analysis with elements of conversation analysis, found not only what the participants talked about in terms of cultural aspects , but also how they talked about cultural issues with regard to negotiating meaning and creating common ground (see section 4.2.2.). It was found that the participants talked about for example stereotypes in terms of food, traditional food, festivities and favorite TV series. A model, showing how the participants mediated and explained cultural concepts and what strategies they applied depending on the level of knowledge of the interlocutor – sharing experience, comparing concepts, explaining a new concept or asking for clarification – has been developed. Interestingly, the participants that explained a (new) cultural concept used a similar discourse pattern that had been further investigated and extracted as discourse structure with discourse elements and examples (expressions). Thus, the analysis brought up many aspects that are evidence-based and frequently used as well as interesting and which are therefore worth teaching.

The **application chapter** used these evidence-based examples and findings from the analysis in the context of EFL teaching. Thirty-eight examples were described in the analysis and application chapter that could be used in the EFL classroom to establish communicative and intercultural competence. Some of these were matched with the requirements in the CEFR for communicative and intercultural competence so as to show to what extent these corpus examples contribute to establishing these competences. These examples meet the requirements in the CEFR and the education standards, which I called the syllabus fit, and are therefore very well suited for application in the EFL classroom. As far as communicative competence is concerned, examples from TaCoCASE were used to improve linguistic competence – teaching vocabulary – and pragmatic competence – teaching use and functions of discourse markers. In terms of communicative strategies, examples from TaCoCASE depicted interaction strategies and mediation strategies that the CEFR demands and that the participants used. The same applies to intercultural competence as well as multimodal features, such as non-verbal communication and online interaction. Selected examples from TaCoCASE illustrated how the

participants used certain strategies and how these examples could be used in the EFL classroom so that the learners identify these strategies and can apply these strategies themselves. Teaching tasks were provided to illustrate what possible tasks in the EFL classroom could look like. In two case studies, an even more in-depth example of a lesson is provided for the following topics: (1) teaching vocabulary and (2) explaining a new cultural concept. A lesson plan and hands-on material, including worksheets, are described in detail so that the full potential of the application is explained and made clear.

Thus, it becomes clear that the application of TaCoCASE in the EFL classroom has much potential and supports in establishing the two main competences in EFL teaching: communicative and intercultural competence. It also became clear that – especially in DDL activities – digital competences can also be established when working with corpora and these main competences can be established all at once, which is very rare in EFL teaching. It is not only the new material – the corpus – that is a rich source and has many benefits; it is also the analysis and application that promotes and new methodology as well as a new mindset that contributes to a modern and authentic way of EFL teaching. It is the combination of providing new material, applying and suggesting new methods and promoting a new mindset that is both the aim of this dissertation and an upgrade for EFL teaching in general.

In providing new material and suggesting new methods, this dissertation is not only a **contribution to the field of EFL teaching**, but also to various other fields. TaCoCASE, the corpus that has especially been designed for this dissertation, provides material and is a rich data source as it is an additional spoken corpus that is available for free. It allows research and application in various fields. Spoken corpora are difficult to compile, and due to these difficulties, written corpora are still more common than spoken corpora. TaCoCASE is an additional spoken corpus that is available for research and thus contributes to the collection of spoken corpora. Much work has gone into compiling the corpus, annotating it and publishing it, which will be well-received by the corpus linguistics community. In fact, conference visits where I presented the corpus suggest great interest in the corpus and some researchers asking for access to TaCoCASE already work with the data. Their research and applications of my corpus will again provide many interesting findings and there is great potential for

groundbreaking research on the basis of the corpus, which enriches research in applied linguistics in general. Thus, TaCoCASE is a valuable contribution to corpus linguistics and applied linguistics.

In addition to the contribution in corpus and applied linguistics, this dissertation suggests many **implications for EFL teaching** in providing new material and new methods, which should lead to a mindset change and thus an upgrade in ELF teaching. The applications of the corpus, including the examples and case studies that were discussed, emphasize the syllabus fit of the corpus and prove that TaCoCASE can be used to improve communicative, intercultural and digital competences in the EFL classroom. The examples and the case studies could be easily be turned into teaching material and integrated in textbooks, which would contribute to more evidence-based and authentic examples in the EFL classroom and in textbooks. New strategies that I found in the analysis and the application of the corpus, should be integrated in the CEFR as well as EFL teaching in general. In looking at discourse markers, non-verbal communication and other multimodal aspects, I propose a more holistic way of teaching languages, including multimodal and pragmatic elements that exist but that have been neglected so far in teaching (Byram 2021). The examples that I provided shed light on these multimodal and pragmatic aspects and suggest that they are worth teaching. Including these aspects depicts a new and more holistic picture of actual language in use, which contributes to establishing authentic communication skills and thus leads to more authentic and competent speakers.

The corpus methods that I applied in the analysis and the way these findings provide information about the selection of words and phrases that are useful to be taught, proposes a new way to more authentic examples in the EFL classroom. Using examples from the corpus and DDL activities as well as these evidence-based findings as a basis for EFL teaching is part of the mindset that I promote in this dissertation and that EFL teaching would benefit from. Policy-makers and textbook designers should follow the suggestions and recommendations in this dissertation and integrate corpora and examples from corpora in EFL teaching. Policy-makers should include using corpora in the syllabi and textbook designers should make use of examples from corpora and DDL activities to bring corpora into the classroom.

The findings and suggestions in this dissertation provide a solid basis for using corpora in EFL teaching and they could be adopted in **other contexts** as well, for example in university teaching, continuing education or lifelong learning as well as in the business context. The material and methods that are presented here, can be used in language classes or intercultural trainings at university or in lifelong learning in a similar manner as discussed here. Communicative as well as (inter)cultural competence can be established through using the corpus and the examples that I presented throughout this work. The findings and examples could be applied as well in a business context, where language courses and intercultural trainings are usually offered for employees. Establishing a good relationship with business partners across the globe is important and English is usually the language of choice (see also the growing research in business ELF – BELF; see for example Ehrenreich 2010; Komori-Glatz 2017). Being aware of language and cultural differences and similarities and how to act in international and intercultural meetings is the key for every employee working in an international environment. Preparing employees for situations like this requires establishing communicative and (inter)cultural competence in workshops and trainings, which can be facilitated with the help of the material and the methods presented in this dissertation. Especially the (inter)cultural examples could serve as critical incidents, which are commonly used in the business context in intercultural trainings (Lüsebrink 2016), but this time in a positive way as the employees could identify and adopt the strategies for mediating and exchanging concepts themselves.

Despite these findings and suggestions, there are some **limitations of this dissertation**. First of all, the corpus designed for the dissertation, TaCoCASE, is quite large for a spoken corpus, but the dataset in general is quite small. It is for example not possible to draw statistical relevant conclusions from this dataset, which would be useful in order to come up with more general hypotheses and suggestions of actual language in use. The quantitative results that I described provide first insights and lead to assumptions that need to be tested, evaluated and proven with a larger dataset. The qualitative analysis generated results, for example, about how discourse markers are used, what functions they have, how the participants talked about cultural concepts in terms of discourse structure and

strategies to mediate cultural concepts, that provide new and interesting findings that could be used in the classroom. However, these findings would probably have to be verified with a larger dataset in order to be representative. The video and audio quality of the data itself could be improved with better technical equipment that exists nowadays. It would probably be easier than to transcribe the data and annotate it, which took quite long in the compilation process and which led to quite a long data collection phase in the whole dissertation project.

In this dissertation, I showed various applications of the corpus when establishing communicative and intercultural competence as demanded by the CEFR in order to prove the syllabus fit. Demonstrating this very broad application leads to a reduction in the depth of the analysis and application, which means that I could not go into very much detail when discussing the results. However, the results provide insights and good starting points for a more in depth analysis in the respective field. Additionally, I could not cover every aspect of competences that the CEFR described and selected the main competences, which means that more work should be done to look at more aspects of the competences, for example socio-linguistic aspects in the examples when establishing linguistic competence. The teaching material and tasks that I presented are also only a selection and much more work and research would have to go into creating even more tasks so that these tasks can be used in the EFL classroom and find their way to the teachers and into the textbooks.

With this dissertation, I have demonstrated that corpora are extremely useful in EFL teaching and to what extent they can be used to establish communicative, intercultural and digital competences. It is clear that this research provides only first insights and more research needs to be conducted. **Further research** would have to be conducted in terms of actually using these examples and suggestions in the EFL classroom and assessing their potential. Some hands-on activities and teaching tasks were provided which could be used in the EFL classroom. It would be nice to see how the learners cope with these tasks and to what extent learners are able to make use of TaCoCASE. It would be interesting to assess the outcome of this suggested methodology, measuring the effects between learners working with corpora and a control group that does not use corpora. Using even more examples from TaCoCASE

and designing tasks around these authentic examples would be very nice, especially for textbook and teaching material designers.

In general, I argue that establishing **intercultural competence** should be given more attention to in the EFL classroom. With technology and translation programs, intercultural competence becomes probably more important than communicative competence. If examples from TaCoCASE were used to teach intercultural competence, this would need to be assessed in some way too. The availability of tools for translation and other tools that make communicative competence seem superfluous, also calls for new assessment methods. Simulations, discussing critical incidents (or just intercultural examples from TaCoCASE) and portfolio work should be integrated even more in the EFL classroom and could substitute written exams or papers. This would also prepare learners for the real world and establish the much needed future skills such as communication, collaboration, creativity and critical thinking. The goal stays the same: preparing learners to act successfully as social agents in intercultural situations (CEFR 2020; KMK 2012). This goal could be better achieved and in a more authentic way with the suggestions in my work. Another main aspect is establishing digital competences. The CEFR does not cover digital competences at all and the KMK has only just started to describe aspects of digital competences in the education standards that need to be integrated into the syllabi in the federal states. Therefore, I could not prove the syllabus fit in terms of digital competences in this work, even though it becomes clear that DDL activities and corpus methods have great potential to establish digital competences (see section 2.1.6.3.; see case studies in section 5.5.)

Thus, the new material and new methodologies that I suggested in this work along with the proposed mindset change, have the potential to modernize and upgrade EFL teaching. It is up to decision makers, policymakers all others responsible in education to follow the suggestions and make this much needed change happen.

REFERENCES

- Aarts, B. 2000. Corpus linguistics, Chomsky and fuzzy tree fragments. In C. Mair & M. Hundt (Eds.), *Corpus Linguistics and Linguistic Theory: Papers from the Twentieth International Conference on English Language Research on Computerized Corpora (ICAME 20)* (pp. 5-13). Amsterdam: Rodopi.
- Abdel-Latif, M. (2020). Corpus literacy instruction in language teacher education: Investigating Arab EFL student teachers' immediate beliefs and long-term practices. *ReCALL*, 33(1), 1–15. <https://doi.org/10.1017/s0958344020000129>.
- Adobe. (2017). *Adobe Premiere Pro* [Computer software]. <https://creativecloud.adobe.com/>.
- Adolphs, S., & Knight, D. (2010). Building a spoken corpus. In A. O'Keeffe, A. & M. McCarthy (Eds) *The Routledge Handbook of Corpus Linguistics* (pp. 38-52). New York: Routledge.
- Agar, M. (1994). *Language shock - Understanding the culture of conversation*. New York: HarperCollins.
- Aijmer, K. (2002). English discourse particles. Evidence from a corpus. *Acta Linguistica Hafniensia*, 36(1), 190-193.
- Allrath, G., Gymnich, M., & Surkamp, C. (2005). Introduction: Towards a Narratology of TV Series. In G. Allrath & M. Gymnich (Eds.), *Narrative strategies in television series*. London: Palgrave Macmillan. https://doi.org/10.1057/9780230501003_1.
- Altenberg, B., & Granger, S. (2001). Grammatical and lexical patterning of *make* in student writing. *Applied Linguistics*, 22(2), 173-194.
- Altmayer, C. (2017). Landeskunde im Globalisierungskontext. Wozu noch Kultur im DaF-Unterricht? In P. Haase & M. Höller (Eds.), *Kulturelles Lernen im DaF/DaZ-Unterricht* (pp. 3-22). Göttingen: Universitätsverlag Göttingen.
- Androutopoulos, J. (2012). Introduction: Language and society in cinematic discourse. *Multilingua*, 31(2-3), 139-154. <https://doi.org/10.1515/multi-2012-0007>.
- Ángel Villegas, J., Jiménez Quintero, J., & Marín Ramirez, L. (2022). *The use of TV series as authentic materials to develop pragmatic competence in EFL contexts: a systematic review*. Universidad Tecnológica de Pereira. <https://hdl.handle.net/11059/14627>.
- Anthony, L. (2021). *AntGram* (version 1.3.0.) [Computer software] <https://www.laurenceanthony.net/software/antgram/>.
- . (2023). *AntConc* (version 4.1.0.) [Computer software] <https://www.laurenceanthony.net/software/antconc/>.
- Antonello, M. (2022). *English as a lingua franca and intercultural communication: communication strategies and meaning negotiation in ELF transcultural contexts*. (Unpublished doctoral dissertation). University of Verona.
- Archer, D., Wilson, A., & Rayson, P. (2002). Introduction to the USAS category system. UCREL. https://ucrel.lancs.ac.uk/usas/usas_guide.pdf (Accessed 27 May 2022).
- Argyle, M. (1988). *Bodily Communication* (2nd edition). London: Routledge.
- Armstrong, M. (2015). Using literature in an EFL context to teach language and culture. *The Journal of Literature in Language Teaching*, 4(2), 7-24.
- Aston, G. (2000). Corpora and language teaching. In L. Burnard & T. McEnery (Eds.), *Rethinking language pedagogy from a corpus perspective: Papers from the Third International Conference on Teaching and Language Corpora* (pp. 7-17). Frankfurt am Main: Peter Lang.
- Baker, P. (2006). *Using corpora in discourse analysis*. London & New York: continuum.
- Baker, W. (2009). The cultures of English as a lingua franca. *TESOL Quarterly*, 43(4), 567–592.
- . (2012). From cultural awareness to intercultural awareness: Culture in ELT. *ELT Journal*, 66(1), 62–70.
- . (2015). *Culture and identity through English as a lingua franca: Rethinking concepts and goals in intercultural communication*. Boston: De Gruyter Mouton.
- Balboni, P. E. (2006). Intercultural communicative competence: a model. *Documents in Language Acquisition and Learning*, 2, 1-56.

- Bangerter, A., & Clark, H. (2003). Navigating joint projects with dialogue. *Cognitive Science* 27, 195-225.
- Barlow, M. (1996). Corpora for theory and practice. *International Journal of Corpus Linguistics*, 1, 1–37.
- Barmeyer, C. I. (2000). *Interkulturelles Management und Lernstile. Studierende und Führungskräfte in Frankreich, Deutschland und Québec*. Frankfurt a. Main: Campus.
- BaseX. (2019). BaseX [Computer software] <https://basex.org/>.
- Baude, O., Blanche-Benveniste, C. & Calas, M. (2010). *Spoken Corpora Good Practice Guide 2006*. Presses Universitaires d'Orléans <https://shs.hal.science/halshs-01165893/file/Spoken%20corpora%20Good%20Practices%20%281%29.pdf> (Accessed 02 December 2023).
- Baumann, Z. (2004). *Identity*. Cambridge: Polity Press.
- Baumgarten, N., & House, J. (2010). *I think and I don't know* in English as lingua franca and native English discourse. *Journal of Pragmatics*, 42(5), 1184-1200.
- Baumgartner, P., & Herber, E. (2013). Höhere Lebensqualität durch interaktive Medien? – Eine kritische Reflexion. *Erziehung und Unterricht*, 163(3-4), 327-335.
- Bednarek, M. (2018). *Language and television series: A linguistic approach to TV dialogue*. Cambridge University Press.
- Beneke, J. (2000). Intercultural competence. In U. Bliesener (Ed.), *Training the Trainers. International Business Communication* (Volume 5, pp. 108-109). Köln: Carl Duisberg Verlag.
- Bennett, G. (2010). Using corpora in the English Language Learning Classroom: Corpus Linguistics for teachers. *ELT journal*. URL: <http://www.international.ucla.edu/media/files/Using-corpora-in-the-language-learning-classroom--Corpus-linguistics-for-teachers-my-atc.pdf> (Accessed 6 January 2023).
- Bennett, M. J. (1993). Towards ethnorelativism: A development model of intercultural sensitivity. In Paige, M. (Ed.), *Education for the Intercultural Experience* (pp. 21-72). Yarmouth: Intercultural Press.
- Bennett, J. M. & Bennett, M. J. (2004). Developing Intercultural Sensitivity. In J. M. Bennett, M. J. Bennett, & D. Landis (Eds), *Handbook of Intercultural Training* (3rd edition, pp. 147-165). Thousand Oaks: Sage.
- Bernardini, S. (2001). "Spoilt for choice": A learner explores general language corpora. In G. Aston (Ed.), *Learning with corpora* (pp. 220-249). Houston: Athelstan.
- . (2004). Corpora in the classroom: an overview and some reflections on future developments. In J. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 5-16). Amsterdam: John Benjamins.
- Bertrand, R., Ferré, G., Blache, P., Espesser, R., & Rauzy, S. (2007). Backchannels revisited from a multimodal perspective. *Auditory-visual Speech Processing* (pp. 1-5). Hilvarenbeek: HAL Open Science. <https://hal.science/hal-00244490> (Accessed 10 June 2024).
- Bhabha, H.K. (1994). *The location of culture*. London: Routledge.
- . (2007): *Die Verortung der Kultur*. Tübingen: Stauffenburg Verlag.
- Biber, D. (2006). *University language - A corpus-based study of spoken and written Registers* [Studies in Corpus Linguistics 23]. Amsterdam: John Benjamins.
- . (2009). A corpus-driven approach to formulaic language in English: Multi-word patterns in speech and writing. *International Journal of Corpus Linguistics*, 14(3), 275-311.
- Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for specific purposes*, 26(3), 263-286.
- Biber, D., Conrad, S., & Cortes, V. (2004). *If you look at. . .*: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25, 371–405.
- Biber, D., Conrad, S. & Leech, G. (2002). *Longman student grammar of spoken and written English*. London: Longman.
- Biber, D., Conrad, S., & Reppen, R. ([1998] 2004). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.

- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. London, UK: Longman.
- Biber, D., & Reppen, R. (2002). What does frequency have to do with grammar teaching? *Studies in Second Language Acquisition*, 24(2), 199-208.
- Biber, D., & Staples, S. (2014). Exploring the prosody of stance: variation in the realization of stance adverbials. In T. Raso & H. Mello (Eds.), *Spoken Corpora and Linguistic Studies* (pp. 271-294). Amsterdam: John Benjamins.
- Birner, B. J. (2013). *Introduction to Pragmatics*. Oxford: Wiley-Blackwell Publishing.
- Björge, A. K. (2010). Conflict or cooperation: The use of backchanneling in ELF negotiations. *English for Specific Purposes*, 29(3), 191-203.
- Blakemore, D. (1988). "So" as a constraint on relevance. In R. Kempson (Ed.), *Mental representations: The interface between language and reality* (pp. 183-195). New York: Cambridge University Press.
- Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain*. New York: McKay.
- BNC. (1994). *British National Corpus* [Corpus]. BNC Consortium. <http://www.natcorp.ox.ac.uk/>
- Bolten, J. (2000). Interkultureller Trainingsbedarf aus der Perspektive der Problemerkahrungen entsandter Führungskräfte. In K. Götz (Ed.), *Interkulturelles Lernen, interkulturelles Training* (pp. 61-80). München: Rainer Hampp.
- Bolden, G. (2006). Little words that matter: Discourse markers "so" and "oh" and the doing of other-attentiveness in social interaction. *Journal of Communication*, 56, 661-688.
- Bolden, G. (2009). Implementing incipient actions: The discourse marker "so" in English conversation. *Journal of Pragmatics*, 41, 974-998.
- Bonnet, A. (2009). Kooperatives Lernen. *Der Fremdsprachliche Unterricht Englisch*, 99, 2 - 8.
- . (2018). Language learners - from learning styles to identity. In C. Surkamp, & B. Viebrock (Eds.), *Teaching English as a Foreign Language* (pp 57-72). Stuttgart: J.B. Metzler.
- Bonnet, A., Decke-Cornill, H., & Hericks, U. (2017). Kooperatives Lernen. In C. Surkamp (Ed.), *Metzler Lexikon Fremdsprachendidaktik: Ansätze – Methoden – Grundbegriffe* (pp. 172-174). Stuttgart: J.B. Metzler.
- Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL*, 21, 37-54.
- Bradner, E., & Mark, G. (2002, November). Why distance matters: effects on cooperation, persuasion and deception. In *Proceedings of the 2002 ACM conference on computer supported cooperative work* (pp. 226-235).
- Breyer, Y. (2009). Learning and teaching with corpora: Reflections by student teachers. *Computer Assisted Language Learning*, 22(2), 153-172. <https://doi.org/10.1080/09588220902778328>
- Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide*. Cambridge: Cambridge University Press.
- Brezina, V., & Pallotti, G. (2019). Morphological complexity in written L2 texts. *Second language research*, 35(1), 99-119.
- Brinton, L. J. (1996). *Pragmatic markers in English – Grammaticalization and discourse functions*. Berlin: De Gruyter.
- Brown, G. & Yule, G. (1983). *Discourse Analysis* [Cambridge Textbooks in Linguistics]. Cambridge: Cambridge University Press.
- Brunner M.-L., & Diemer, S. (2020). "Tell me about food and I tell you who you are"- Expert identity in intercultural food discourse via Skype. In S. Rüdiger, & S. Mühleisen (Eds.), *Talking about Food. The social and the global in eating communities* (pp. 167-187). Amsterdam: John Benjamins. <https://doi.org/10.1075/impact.47.09die>.
- . (2018). "You Are Struggling Forwards, and You Don't Know, and Then You ... You do code-switching..." — Code-switching in ELF Skype conversations. *Journal of English as a Lingua Franca*, 7(1), 59-88.
- Brunner, M.-L. (2021). *Understanding intercultural communication - Negotiating meaning and identities in English as a lingua franca Skype conversations*. Berlin: J.B. Metzler.

- . (2015). *Negotiating conversation starts in the corpus of academic spoken English*, MA Thesis, Saarland University, Saarbrücken.
- Burns, A. (2001). Analysing spoken discourse: implications for TESOL. In A. Burns, & C. Coffin, (Eds.) *Analysing English in a global context : A reader* (pp. 123-148). London: Routledge.
- Buyse, L. (2012). So as multifunctional discourse marker in native and learner speech. *Journal of Pragmatics*, 44, 1764-1782.
- Byram, M. (1989). Intercultural education and foreign language teaching. *World Studies Journal*, 7(2), 4-7.
- . (1991). Teaching language and culture: towards an integrated model. In D. Buttjes & M. Byram (Eds.), *Mediating languages and cultures: Towards an intercultural theory of foreign language education* (pp. 17-30). Clevedon: Multilingual Matters.
- . (1995). Intercultural competence and mobility in multinational context.: a European view. In M. L. Tickoo (Ed.), *Language and culture in multilingual societies: Viewpoints and visions* (pp. 21-36). Singapore: SEAMEO Regional Language Centre.
- . (1997). *Teaching and Assessing Intercultural Communicative Competence*. Clevedon: Multilingual Matters.
- . (2008). *From foreign language education to education of intercultural citizenship: Essays and reflections*. Clevedon: Multilingual Matters.
- . (2012). Conceptualizing intercultural (communicative) competence and intercultural citizenship. In J. Jackson (Ed.), *The Routledge Handbook of Language and Intercultural Communication* (pp.85-99). London & New York: Routledge.
- . (2021). *Teaching and assessing intercultural communicative competence: Revisited* (2nd edition). Bristol: Multilingual matters.
- Byram, M. & Fleming, M. (Eds.). (1998). *Language learning in intercultural perspective: Approaches through drama and ethnography*. Cambridge: Cambridge University Press.
- Caines, A., McCarthy, M. J. & O’Keeffe, A. (2016). Spoken language corpora and pedagogic applications. In F. Farr & L. Murray (Eds.), *Routledge Handbook of Language Learning and Technology* (pp. 348-361). London: Routledge.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47.
- CANCODE (1994 - 2004). *Cambridge and Nottingham Corpus of Discourse in English* [Corpus]. University of Nottingham. <http://www.cambridge.org/elt/corpus/cancode.html>.
- Canva. (2024). *Canva* [Design software]. https://www.canva.com/de_de/.
- Carter, R. A. (1998). Orders of reality: CANCODE, communication and culture. *ELT journal*, 52, 43-56.
- Carter, R. A. & McCarthy, M. (1995). Grammar and the spoken language. *Applied Linguistics*, 16(2), 141–158.
- CASE. (2012-2023). *The CASE project* [Corpus]. Diemer, S., Brunner, M.-L., Collet, C., & Schmidt, S. Saarland University & Trier University of Applied Sciences. <http://umwelt-campus.de/case>.
- Celce-Murcia, M., Dörnyei, Z. & Thurrell, S. (1995). Communicative competence: A pedagogically motivated model with content specifications. *Issues in Applied Linguistics*, 6 (2), 5–35.
- Chafe, W. (2007). *The importance of not being earnest: the feeling behind laughter and humor*. Amsterdam: John Benjamins.
- Chambers, A. (2019). Towards the corpus revolution? Bridging the research-practice gap. *Language Teaching*, 52(4), 460–475. <https://doi.org/10.1017/S0261444819000089>.
- Chatman, S. (1999). New directions in voice-narrated cinema’. In D. Herman (Ed.), *New perspectives on narrative analysis* (pp. 315–39). Ohio: Ohio State University Press.
- Chen, C. W.-Y. (2006a). The use of conjunctive adverbials in the academic papers of advanced Taiwanese EFL learners. *International Journal of Corpus Linguistics*, 11(1), 113–130.
- Chen, L. L. (2006b). *Writing Chinese: Reshaping Chinese cultural identity*. Gordonsville: Palgrave Macmillan.
- Cheng, W. (2010). What can a corpus can tell us about language teaching? In A. O’Keeffe, & M. McCarthy (Eds.), *The Routledge Handbook of Corpora and English Language Teaching and Learning* (pp. 319-332). New York: Routledge.

- Cheng, W., Warren, M., & Xun-Feng, X. (2003). The language learner as language researcher: Putting corpus linguistics on the timetable. *System*, 31(2), 173-186.
- Christoffersen, K. O. (2017). Comparing native speaker ratings and quantitative measures of oral proficiency in IELTS interviews. *Elia*, 17, 233-250. <http://dx.doi.org/10.12795/elia.2017.i17.10>.
- Cobb, T. (2002). *Vocabprofile English* [Computer software]. <https://www.lexutor.ca/vp/eng/> (Accessed 10 March 2017).
- Cobb, T., & Horst, M. (2015). Learner corpora and lexis. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge Handbook of Learner Corpus Research* (pp. 185-206). Cambridge: Cambridge University Press.
- Cogo, A., & Dewey, M. (2012). *Analyzing English as a Lingua Franca: a corpus-driven Investigation*. London; New York: Continuum.
- Cohn, R., & Terfurth, C. (1997). *Lebendiges Lehren und Lernen. TZI macht Schule*. Stuttgart: Klett-Cotta.
- Collet, C., & Ankerstein, C. (2015). A corpus-based exploration of connector usage in native and non-native English writing. *Saarland Working Papers in Linguistics*, 5, 1-7. <http://doi.org/10.22028/D291-23644>.
- Collet, C., & Diemer, S. (2014). Transcribing academic spoken English - challenges and issues" [Conference paper]. *12th European Society for the Study of English Conference (ESSE 12)*, Košice, Slovakia.
- Collet C., Diemer S., Brunner M.-L. (2021). So in video-mediated communication in the Expanding Circle. *World Englishes*, 40, 594-610. <https://doi.org/10.1111/weng.12543>.
- Cook, V. (2008). *Second language learning and teaching*. London: Hodder Education.
- . (1999). Going beyond the native speaker in language teaching. *TESOL Quarterly*, 33 (2), 185-209. <https://doi.org/10.2307/3587717>.
- Coperias Aguilar, M. (2002). Intercultural communicative competence: A step beyond communicative competence. *Estudios de lingüística inglesa aplicada*, 3, 85-102.
- Corbett, J. (2003). *An Intercultural approach to English language teaching*. Clevedon: Multilingual Matters.
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23, 397-423.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. <https://www.coe.int/en/web/common-european-framework-reference-languages> (Accessed 12 March 2023).
- . (2012). *Framework of reference for pluralistic approaches to languages and cultures - competences and resources*. <https://www.ecml.at/Resources/ECMLresources/tabid/277/ID/20/language/en-GB/Default.aspx> (Accessed 14 March 2023).
- . (2016). *Guide for the development and implementation of curricula for plurilingual and intercultural education*. <https://www.coe.int/en/web/language-policy/guide-for-the-development-and-implementation-of-curricula-for-plurilingual-and-intercultural-education> (Accessed 14 March 2023).
- . (2018). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Companion Volume with new descriptors*. <https://www.coe.int/en/web/common-european-framework-reference-languages> (Accessed 12 March 2023).
- . (2020). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Companion Volume*. <https://www.coe.int/en/web/common-european-framework-reference-languages> (Accessed 12 March 2023).
- . (2022). *Enriching 21st century Language Education - the CEFR Companion volume in practice*. https://www.coe.int/en/web/common-european-framework-reference-languages/home/-/asset_publisher/3omt7fb9wnHZ/content/enriching-21st-century-language-education-the-cefr-companion-volume-in-practice (Accessed 12 February 2023)
- Counihan, C. (1999). *The anthropology of food and body: Gender, meaning and power*. New York: Routledge.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34, 213-38.

- Crystal, D. (2003 [2012]). *English as a Global language* (2nd edition). Cambridge: Cambridge University Press.
- Darvin, R. & Norton, B. (2015). Identity and a model of investment in applied linguistics. *Annual Review of Applied Linguistics*, 35, 36–56.
- Deardorff, D. K. (2004). The identification and assessment of intercultural competence as a student outcome of internationalization at institutions of higher education in the United States (Doctoral dissertation). North Carolina State University. <https://repository.lib.ncsu.edu/server/api/core/bitstreams/c047bd17-d120-4942-9a77-f043746e5097/content> (Accessed 17 March 2023).
- Deardorff, D. K. (2006). Identification and assessment of intercultural competence as a student outcome of internationalization. *Journal of Studies in International Education*, 10 (3), 241-266.
- Deci, E. L., & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39(2), 223-238.
- Decke-Cornill, H. & Küster, L. (2015). *Fremdsprachendidaktik: Eine Einführung*. Tübingen: Narr Francke Attempto Verlag.
- DeepL. (2024). *DeepL Translator* [Computer software]. DeepL SE. <https://www.deepl.com/de/translator>.
- Dervin, F. (2012). Cultural identity, representation and othering. In J. Jackson (Ed.), *The Routledge Handbook of Language and Intercultural Communication* (pp. 181-194). London & New York: Routledge.
- Diani, G. (2008). Emphasizers in spoken and written academic discourse: The case of *really*. *International Journal of Corpus Linguistics*, 13(3) 296-321.
- Diemer, S., & Brunner, M. L. (2020). Tell me about food and I tell you who you are. In S. Rüdiger & S. Mühleisen (Eds.), *Talking about food: The social and the global in eating communities* (pp. 167-187). Amsterdam: John Benjamins.
- Diemer, S., Brunner, M.-L., & Schmidt, S. (2016). Compiling computer-mediated spoken language corpora – Key issues and recommendations. *International Journal of Corpus Linguistics*, 21 (3), 348-371.
- Dimoski, B., Yujobo, Y. J. & Imai, M. (2016). Exploring the effectiveness of communication strategies through pro-active listening in ELF-informed pedagogy. *Language Education in Asia* 7 (2), 67–87.
- Doff, S. (2018). English language teaching and English language education - history and methods. In C. Surkamp & B. Viebrock (Eds.), *Teaching English as a Foreign Language* (pp. 1-16). Stuttgart: Metzler.
- Döbeli Honegger, B. (2021). Was machen wir mit der Digitalisierung?. *Pädagogik*, 5(21), 41-46. <https://beat.doebe.li/publications/2021-doebeli-honegger-was-machen-wir-mit-der-digitalisierung.pdf> (Accessed 4 January 2024).
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.
- .(1995). On the teachability of communication strategies. *TESOL Quarterly*, 29(1), 55-85.
- Dörnyei, Z. and J. Kormos (1998), 'Problem-Solving Mechanisms in L2 Communication: A Psycholinguistic Perspective', *Studies in Second Language Acquisition*, 20 (3): 349-385.
- Dreschler, M. (2020). The moral taste of food: A discourse analysis of social media discussions about vegetarianism and veganism. In S. Rüdiger & S. Mühleisen (Eds.), *Talking about food: The social and the global in eating communities* (pp. 35-56). Amsterdam: John Benjamins.
- Dressler, R. A., & Kreuz, R.J. (2000). Transcribing oral discourse: A survey and a model system. *Discourse Processes*, 29(1), 25-36.
- Edmondson, W., & House, J. (1981). *Let's talk and talk about it: A pedagogic interactional grammar of English*. München: Urban and Schwarzenberg.
- Edwards, J. A. (1993). Principles and contrasting systems of discourse transcription. In J. A. Edwards & M. D. Lampert (Eds.), *Talking data: Transcription and coding in discourse research* (pp. 3-32). Hillsdale: Lawrence Erlbaum Associates.
- . (2009). *Language and identity*. Cambridge: Cambridge University Press.

- Ehrenreich, S. (2010). English as a business lingua franca in a German multinational corporation: Meeting the challenge. *The Journal of Business Communication*, 47(4), 408-431.
- Ehrhart, S. (2019). Staatliche (kollektive) und individuelle Mehrsprachigkeit. In C. Fäcke & F.-J. Meißner (Eds.), *Handbuch Mehrsprachigkeits- und Mehrkulturalitätsdidaktik* (pp. 25-29). Tübingen: Narr Francke Attempto.
- ELFA. (2008). *The Corpus of English as a Lingua Franca in Academic Settings* [Corpus]. Mäuranen, A. University of Helsinki. <http://www.helsinki.fi/elfa/elfacorpus>.
- Ellis, N. C. (2008). Phraseology: The periphery and the heart of language. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 1–13). Amsterdam: John Benjamins.
- Ellis, R. (1997). *Second language acquisition*. Oxford: Oxford University Press.
- Elsner, D. (2018). Institutionalised foreign language teaching - teaching English at different levels. In C. Surkamp & B. Viebrock (Eds.), *Teaching English as a foreign language* (pp. 17-38). Stuttgart: Metzler.
- Erll, A. & Gymnich, M. (2021). *Interkulturelle Kompetenzen* (Volume 6). Stuttgart: Klett.
- Erpenbeck, J. (2012). Was "sind" Kompetenzen? In W. Faix (Ed.), *Kompetenz. Festschrift Prof. Dr. John Erpenbeck zum 70. Geburtstag* (pp. 3-59). Stuttgart: Steinbeis-Edition.
- European Commission. (2017). *DigiCompEdu: European Framework for the Digital Competence of Educators*. <https://publications.jrc.ec.europa.eu/repository/handle/JRC107466> (Accessed 14 March 2023).
- Evison, J. (2010). What are the basics of analysing a corpus? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 122-135). New Work: Routledge.
- Faerch, C., & Kasper, G. (1984). Two ways of defining communication strategies. *Language Learning*, 34(1), 45-63.
- Fahey Palma, T. (2022). Corpora for teaching culture and intercultural communication. In R.R. Jablonkai, & E. Csomay (Eds.), *The Routledge Handbook of Corpora and English Language Teaching and Learning* (pp. 116-130). New York: Routledge.
- Fairclough, N. (1995). *Critical discourse analysis: The critical study of language*. London & New York: Longman.
- Fantini, A. (2019). *Intercultural communicative competence in educational exchange: a multinational perspective*. London: Routledge.
- . (2000). A central concern: Developing intercultural communicative competence. *School for International Training Occasional Papers Series*, 1, 25-42.
- Fantini, A. and Tirmizi, A. (2006). Exploring and assessing intercultural competence. *World Learning Publications*, 1. http://digitalcollections.sit.edu/worldlearning_publications/1 (Accessed 13 April 2022).
- Farr, F., & Leńko-Szymańska, A. (2023). Corpora in English language teacher education: Research, integration, and resources. *TESOL Quarterly*, 58 (3), 1181-1192.
- Fellner, A. M., Hamscha, S., Moos, J. & Heissenberger, K. (2014). Introduction: Is American culture popular because it's cool? – or is it cool because it's popular? In A. M. Fellner, S. Hamscha, J. Moos & Heissenberger K. (Eds.), *Is it 'cause it's cool? Affective encounters with American culture* (pp. 9-24). Vienna: LIT.
- Fellner, A. M., Ostapchuk, T., & Schlimbach, B. (2018). Introduction: (Pop)Cultures on the move. In A. M. Fellner, T. Ostapchuk, & B. Schlimbach (Eds.), *(Pop) Cultures on the move: transnational identifications and cultural exchange between east and west* (pp. 7-16). Saarbrücken: Saarland University Press.
- Firth, J.R. (1957). A synopsis of linguistic theory 1930-1955. *Studies in Linguistic Analysis: Special volume of the Philological Society* (pp. 1 - 32). Oxford: Blackwell.
- Firth, A. (2009). The lingua franca factor. *Intercultural pragmatics*, 6(2), 147-170.

- Fischer, K. (2006). Frames, constructions, and invariant meaning. The functional polysemy of discourse particles. In K. Fischer (Ed.), *Approaches to Discourse Particles* (pp. 427-447). Amsterdam: Elsevier.
- Fischler, C. (1988). Food, self and identity. *Social Science Information*, 27(2), 275-292.
- Fish, R. S., Kraut, R. E., Root, R. W., & Rice, R. E. (1993). Video as a technology for informal communication. *Communications of the ACM*, 36(1), 48-61.
- Fligelstone, S. (1993). Some reflections on the question of teaching, from a corpus linguistics perspective. *ICAME Journal*, 17, 97-109.
- Fox, G. (1987). The case for examples. In J. M. Sinclair (Ed.), *Looking up: An account of the COBUILD project in lexical computing* (pp. 137-149). London: Collins ELT.
- Fraser, B. (1990). An approach to discourse markers. *Journal of Pragmatics*, 14(3), 383-398.
- . (1999). What are discourse markers? *Journal of Pragmatics*, 31, 931-952.
- Freie und Hansestadt Hamburg - Behörde für Schule und Berufsbildung. (2022). *Bildungsplan Studienstufe Neuere Fremdsprachen Englisch*. <https://www.hamburg.de/contentblob/16762884/d0738e7bafad8496f4ef1a6b9cefb926/data/fsp-englisch-gyo-2022.pdf> (Accessed 12 December 2023)
- Friedmann, J. (1994). *Cultural identity and global process*. London: Sage.
- Fung, L., & Carter, R. (2007). Discourse markers and spoken English: Native and learner use in pedagogic settings. *Applied Linguistics*, 28(3), 410-439. DOI:10.1093/applin/amm030.
- Galloway, N. (2018). ELF and ELT teaching materials. In J. Jenkins, W. Baker, & M. Dewey (Eds.), *The Routledge Handbook of English as a Lingua Franca* (pp. 468-480). New York: Routledge.
- Gee, J. P. (2007). *Social linguistics and literacies: Ideology in discourses* (3rd edition). New York: Routledge.
- . (2011). *An introduction to discourse analysis: Theory and method* (3rd edition). New York: Routledge.
- Gee, M. (2019). XTranscript [Computer software]. <http://mattgee.net/xtranscript/convert.cgi>
- Geluykens, R. (1993). Topic introduction in English conversation. *Transactions of the Philological Society*, 91(2), 181-214.
- Germov, J., & Williams, L. (Eds.). (2008). *A Sociology of food and nutrition: The social appetite*. Oxford: Oxford University Press.
- Gesellschaft für Informatik. (2016). *Dagstuhl-Erklärung*. https://gi.de/fileadmin/GI/Hauptseite/Themen/Dagstuhl-Erklärung_2016-03-23.pdf (Accessed 11 February 2023)
- Gilmore, A. (2004). A comparison of textbook and authentic interactions. *ELT Journal*, 58(4), 363-374.
- . (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching*, 40, 97-118.
- Gilquin, G., & Granger, S. (2010). How can data-driven learning be used in language learning? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 359-370). New York: Routledge.
- Girard, M. & Sionis, C. (2004). The functions of formulaic speech in the L2 class. *Pragmatics*, 14(1), 31-53.
- Godley, A. J. (2012). Intercultural discourse and communication in education. In C. B. Paulston, S. F. Kiesling, E. S. Rangel (Eds.), *The Handbook of Intercultural Discourse and Communication* (pp. 449-481). Oxford: Blackwell.
- Goertler, S. (2009). Using computer-mediated communication (CMC) in language teaching. *Die Unterrichtspraxis/Teaching German*, 42(1), 74-84.
- Göbel, K. (2011). Interkulturelle Kompetenz und Englischunterricht. In: C. Allemann-Ghionda, & W.-D. Bukow (Eds.), *Orte der Diversität* (pp. 191-208). Wiesbaden: VS Springer.
- Göbel, K., & Hesse, H.-G. (2004). Vermittlung interkultureller Kompetenz im Englischunterricht. Eine curriculare Perspektive. *Zeitschrift für Pädagogik*, 50(6), 818-834.
- Götz, S., & Schilk, M. (2011). Formulaic sequences in spoken ENL, ESL and EFL. Focus on British English, Indian English and learner English of advanced German learners. In M. Hundt & J. Mukherjee

- (Eds.), *Exploring second-language varieties of English and learner Englishes: Bridging a paradigm gap* (pp. 79-100). Amsterdam: John Benjamins.
- Gohard-Radenkovic, A., Lussier, D., Penz, H., & Zarate, G. (2004). Chapter 2: Reference fields and methodologies. In: G. Zarate (Ed.) *Cultural mediation in language learning and teaching*. Strasbourg: Council of Europe Publishing.
- Gong, W. (2005). English in computer-mediated environments: A neglected dimension in large English corpus compilation. *Proceedings of Corpus linguistics conference series 1* (1).
- Gonzalez, M. (2013). *The intricate relationship between measures of vocabulary size and lexical diversity as evidenced in non-native and native speaker academic compositions* (Doctoral dissertation). College of Education and Human Performance: University of Central Florida.
- Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics*, 32(10), 1489–1522.
- Goodwin, M. H., & Goodwin, C. (2000). Emotion within situated activity. Communication: An arena of development. In N. Budwig, I. C. Uzgriz & J. V. Wertsch (Eds.), *Communication* (pp. 33-53). Stamford: Ablex.
- Graddol, D. (1998). *The future of English?* London: The British Council.
- Granath, S. (2009). Who benefits from learning how to use corpora? In K. Aijmer (Ed.), *Corpora and language teaching* (pp. 47–65). Amsterdam: John Benjamins.
- Gries, S. T. (2016). *Quantitative corpus linguistics with R: A practical introduction* (2nd edition). London: Routledge.
- Gullberg, M. (1998). Gesture as a communication strategy in second language discourse: A study of learners of French and Swedish. *Travaux de l'institut de linguistique de Lund 35*. Lund: Lund University.
- Gumperz, J. J. (1982). *Discourse Strategies*. Cambridge: Cambridge University Press
- Hadley, G. (2002). An introduction to data-driven learning. *RELC Journal*, 33(2), 99-124.
- Hall, S. (1992). The question of cultural identity. In S. Hall, D. Held & A. G. McGrew (Eds.), *Modernity and its Futures* (pp. 274–316). Oxford: Polity Press.
- Hall, E. T. (1959). *The silent language*. New York: Anchor Books, Doubleday.
- . (1966). *The hidden dimension*. New York: Anchor Books, Doubleday.
- Hall, E. T. & Hall, M. R. (1990). *Understanding cultural differences. Germans, French and Americans*. Yarmouth: Intercultural Press.
- Hall, S. ([1997] 2001). The work of representation. In S. Hall (Ed.), *Representation: Cultural representations and signifying practices* (pp. 41–64). London: Sage.
- Hallet, W. (2008). Diskursfähigkeit heute. Der Diskursbegriff in Piephos Theorie der kommunikativen Kompetenz und seine zeitgemäße Weiterentwicklung für die Fremdsprachendidaktik. In M. Legutke (Ed.), *Kommunikative Kompetenz als fremdsprachendidaktische Vision* (pp. 76-96). Tübingen: Gunter Narr Verlag.
- Halliday, M. A. K. (1985). *Spoken and written language*. Geelong Victoria: Deakin University.
- Halliday, M. A. K. (1973). *Explorations in the Functions of Language*. London: Edward Arnold.
- Harwood, N. (2005). 'We do not seem to have a theory... The theory I present here attempts to fill this gap'. Inclusive and exclusive pronouns in academic writing. *Applied linguistics*, 26(3), 343-375.
- Haß, F. (Ed.). (2006). *Fachdidaktik Englisch: Tradition. Innovation. Praxis*. Stuttgart: Ernst Klett Verlag.
- Haß, F. (2009). *Kooperatives Lernen mit Green Line*. Ernst Klett Verlag.
https://www2.klett.de/sixcms/media.php/229/W_KV_Kooperative_Lernformen_Uebersicht.pdf (Accessed 12 May 2023).
- Hellerman, J., & Vergun, A. (2007). Language which is not taught: The discourse markers use of beginning adult learners of English. *Journal of Pragmatics*, 39(1), 157-179.
- Heringer, H.-J. (2004): *Interkulturelle Kommunikation. Grundlagen und Konzepte*. Tübingen/Basel: Francke.
- Herring, S. C. (2001). Computer-mediated discourse. In D. Schiffrin, D. Tannen, & H. Hamilton (Eds.), *The Handbook of discourse analysis* (pp. 612–634). Oxford: Blackwell.
- . (2002). Computer-mediated communication on the internet. *Annual Review of Information Science and Technology*, 36, 109–168.

- Hesse, H.G. & Göbel, K. (2007). Interkulturelle Kompetenz. In B. Beck, & E. Klieme (Eds.), *Sprachliche Kompetenzen - Konzepte und Messung - DESI-Studie* (pp. 253-269). Weinheim: Beltz.
- Hessisches Kultusministerium. (2022). *Kerncurriculum gymnasiale Oberstufe Englisch*. https://kultusministerium.hessen.de/sites/kultusministerium.hessen.de/files/2022-09/kcgo_englisch_stand_august_2022_0.pdf (Accessed 12 December 2023).
- Hoey, M. (1991). *Patterns of lexis in text*. Oxford: Oxford University Press.
- . (2005). *Lexical priming: A new theory of words and language*. London: Routledge.
- Hofstede, G. ([1993] 2013). *Interkulturelle Zusammenarbeit: Kulturen – Organisationen – Management*. Wiesbaden: Springer-Verlag.
- Holliday, A. (2020). Culture, communication, context, and power. In J. Jackson (Ed.), *The Routledge Handbook of Language and Intercultural communication* (2nd edition, pp. 39-54). New York: Routledge.
- Holme, R. (2009). *Cognitive linguistics and language teaching*. London: Palgrave.
- House, J. (2013). Developing pragmatic competence in English as a lingua franca: Using discourse markers to express (inter)subjectivity and connectivity. *Journal of Pragmatics*, 59, 57-67.
- Howarth, C. (2002). Identity in whose eyes? The role of representations in identity construction. *Journal for the Theory of Social Behaviour*, 32(2), 145-162.
- Howe, M. L. (1991). *Topic change in conversation* (Unpublished doctoral dissertation). Lawrence: University of Kansas.
- Hua, Z. (2015). Negotiation as the way of engagement in intercultural and lingua franca communication: Frames of reference and interculturality. *Journal of English as a Lingua Franca*, 4(1), 63–90.
- Hua, Z. (2016). Identifying Research Paradigms. In Z. Hua (Ed.), *Research Methods in Intercultural Communication: A Practical Guide* (pp. 3-22). Oxford: Wiley-Blackwell Publishing.
- Hundt, M., & Mukherjee, J. (2011). Introduction: Bridging a Paradigm Gap. In J. Mukherjee & M. Hundt (Eds.), *Exploring Second-Language Varieties of English and Learner Englishes: Bridging a Paradigm Gap [Studies in Corpus Linguistics 44.]* (pp. 1–6). Amsterdam: John Benjamins.
- Hüllen, W. (2005): *Kleine Geschichte des Fremdsprachenlernens*. Berlin: Erich Schmidt Verlag.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Hunston, S., & Francis, G. (2000). *Pattern grammar - a corpus-driven approach to the lexical grammar of English*. Amsterdam: John Benjamins.
- Hutchby, I., & Wooffitt, R. (2008). *Conversation Analysis* (2nd edition). Cambridge: Polity Press.
- Hutz, M. (2018). Focus on form - the lexico-grammar approach. In C. Surkamp & B. Viebrock (Eds.), *Teaching English as a Foreign Language* (pp. 133-158). Stuttgart: Metzler.
- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for specific purposes*, 27(1), 4-21.
- Hymes, D. (1972). *On communicative competence*. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269–293). Harmondsworth: Penguin.
- ICE. (1990- 2024). *International Corpus of English [Corpus]*. Hundt, M.; Lehmann, M. & Schneider, G. University of Zurich. <https://www.ice-corpora.uzh.ch/en.html>
- Jablonkai, R. R., & Csomay, E. (Eds.). (2023). *The Routledge handbook of corpora and English language teaching and learning*. London: Routledge.
- Jefferson, G., Sacks, H., & Schegloff, E. A. (1987). Notes on laughter in the pursuit of intimacy. In G. Button & J. R. E. Lee (Eds.), *Talk and Social Organization* (pp. 152-205). Clevedon: Multilingual Matters.
- Jenkins, J. (2015a). *Global Englishes: A resource book for students* (3rd edition), London: Routledge.
- . (2015b). Repositioning English and Multilingualism in English as a Lingua Franca. *Englishes in Practice*, 2(3), 49–85.
- Jenkins, J., Baker, W., & Dewey, M. (Eds.) (2018). *The Routledge Handbook of English as a Lingua Franca*. Milton Park: Routledge.
- Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: A developmental perspective. *Working papers/Lund University, Department of Linguistics and Phonetics*, 53, 61-79.

- Johns, T., & King, P. (Eds.). (1991). *Classroom Concordancing [Special issue of ELR Journal]* (Volume 4). Birmingham: Birmingham University Press.
- Johns, T. (1997). Contexts: The background, development and trialling of concordance-based CALL program. In A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles (Eds.), *Teaching and language corpora* (pp. 100-115). London: Longman.
- Johns, T. (2002). Data-driven learning: the perpetual challenge. In B. Kettmann & G. Marko (Eds.), *Teaching and Learning by doing corpus analysis* (pp. 107-117). Amsterdam: Rodopi.
- Johnson, A. (2002). "So. . .?": Pragmatic implications of so-prefaced questions in formal police interviews. In J. Cotterill (Ed.), *Language in the Legal Process* (pp. 91-110). New York: Palgrave Macmillan.
- Johnson, K. (2008). *An Introduction to foreign language learning and teaching*. London: Routledge.
- Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk & H. Widdowson (Eds.), *English in the World: Teaching and learning the language and literatures* (pp. 11-30). Cambridge: Cambridge University Press.
- . (1992). *The Other Tongue: English across cultures*. Illinois: University of Illinois Press.
- Kasper, G. (2006). Beyond repair: Conversation analysis as an approach to SLA. *AILA Review*, 19, 83-99.
- Kaur, J. (2009). Pre-empting problems of understanding in English as a lingua franca. In A. Mauranen & E. Ranta (Eds.), *English as a Lingua Franca: Studies and Findings*, 107, Newcastle upon Tyne: Cambridge Scholars Publishing
- (2010). Achieving Mutual Understanding in World Englishes. *World Englishes*, 29(2), 192-208.
- Kennedy, C., & Miceli, T. (2001). An evaluation of intermediate students' approaches to corpus investigation. *Language Learning and Technology*, 5(3), 77-90.
- Kennedy, G. (1992). Preferred ways of putting things with implications for language teaching. In J. Svartvik (Ed.), *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82 Stockholm, 4-8 August 1991* (pp. 335-373). Berlin: Mouton de Gruyter.
- Kettunen, K. (2014). Can type-token ratio be used to show morphological complexity of languages? *Journal of Quantitative Linguistics*, 21(3), 223-245.
- Kim, J. Y. (2014). Predicting L2 Writing Proficiency Using Linguistic Complexity Measures: A Corpus-Based Study. *English teaching*, 69(4), 27-51.
- Kirchhoff, P. (2018). Productive competences - speaking, writing, mediating. In C. Surkamp & B. Viebrock (Eds.), *Teaching English as a Foreign Language* (pp. 109-132). Stuttgart: Metzler.
- Kirkpatrick, A. (2007). The communicative strategies of ASEAN speakers of English as a lingua franca. In D. Prescott (Ed.), *English in Southeast Asia: Varieties, literacies and literatures* (pp. 118-137). Newcastle: Cambridge Scholars Press.
- Kultusministerkonferenz (Eds.). (2012). *Bildungsstandards für die fortgeführte Fremdsprache (Englisch / Französisch) für die Allgemeine Hochschulreife*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2012/2012_10_18-Bildungsstandards-Fortgef-FS-Abi.pdf (Accessed 12 February 2019).
- . 2016. *Bildung in der digitalen Welt*. https://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2018/Digitalstrategie_2017_mit_Weiterbildung.pdf (Accessed 14 March 2023).
- . 2021. *Lehren und Lernen in der digitalen Welt. Die ergänzende Empfehlung zur Strategie "Bildung in der digitalen Welt"*. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2021/2021_12_09-Lehren-und-Lernen-Digi.pdf (Accessed 14 March 2023).
- Koehler, M., Mishra, P., Kereluik, K., Shin, T. S., & Graham, C. R. (2014). The technological pedagogical content knowledge framework. In J. Spector, J. M. Merrill, J. Elen, M. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 101-111). New York: Springer.
- Koester, A. (2022). Why face-to-face communication matters: A comparison of face-to-face and computer-mediated communication. In F. Rosette-Crake & E. Buckwalter (Eds.), *COVID-19, communication and culture* (pp. 115-134). London: Routledge.
- Komori-Glatz, M. 2017. (B)ELF in multicultural student teamwork. *JELF*, 6(1), 83-109.

- Koops, C., & Lohmann, A. (2016). Discourse marker sequencing and grammaticalization. In G. Kaltenböck, E. Keizer, & A. Lohmann (Eds.), *Outside the Clause: Form and function of extra-clausal constituents [Studies in Language Companion Series 178]* (pp. 417-446). Amsterdam: John Benjamins.
- König, L., Schädlich, B., & Surkamp, C. (2022). *unterricht_kultur_theorie: Kulturelles Lernen im Fremdsprachenunterricht gemeinsam anders*. Berlin: J.B. Metzler.
- Kramer, J., Moeller, A. & Nugent, K. (2014). Building intercultural competence in the language classroom. Faculty publications: Department of Teaching, Learning and Teacher Education, 16, 1-16.
- Kramsch, C. (1991). Culture in language learning: A view from the United States. In K. de Bot, R. Ginsberg & C. Kramsch (Eds.), *Foreign language research in a cross cultural perspective* (pp. 217-240). Philadelphia: John Benjamins.
- . (1998a). *Language and Culture*. Oxford: Oxford University Press.
- . (1998b). The privilege of the intercultural speaker. In M. Byram & M. Fleming (Eds.), *Language Learning in Intercultural Perspective: Approaches through Drama and Ethnography* (pp. 16–31). Cambridge: Cambridge University Press.
- Kramsch, C., & Hua, Z. (2016). Language and culture in ELT. In G. Hall (Ed.), *The Routledge Handbook of English Language Teaching* (pp. 38-50). London: Routledge.
- Krashen, S., & Terrell, T. (1988). The natural approach. Language acquisition in the classroom. In S. Krashen (Ed.), *Second language acquisition and second language learning - principles and practices in second language acquisition* (pp. 23-51). Hemel Hempstead: Prentice Hall.
- Krommer, A. (2018). *Wider den Mehrwert! Oder: Argumente gegen einen überflüssigen Begriff*. <https://axelkrommer.com/2018/09/05/wider-den-mehrwert-oder-argumente-gegen-einen-ueberfluessigen-begriff/> [Accessed 12 March 2023]
- Kung, F. W. (2013). Fostering EFL Students' L2 Listening and Speaking Competence through TV Series-Based Instruction: Methods and Applications. *Studies in English Language and Literature*, 32, 113-141.
- Ladel, S. (2018). Sinnvolle Kombination virtueller und physischer Materialien. In S. Ladel, J. Knopf, A. Weinberger (Eds.), *Digitalisierung und Bildung* (pp. 3 - 22). Wiesbaden: Springer VS.
- Laitinen, M. (2015). From diachrony to typology: testing typological profiling on advanced non-native data [Conference paper]. *Changing English 2015: Integrating cognitive, social & typological perspectives, 8-10 June*, Helsinki.
- Lam, P. W. (2010). Toward a functional framework for discourse particles: A comparison of *well* and *so*. *Text & Talk*, 30, 657-677.
- Larrea Espinar, A. M., & Raigón Rodríguez, A. R. (2019). Sitcoms as a tool for cultural learning in the EFL classroom. *Pixel-Bit: Revista de Medios y Educación*, 56, 33-50.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307–322.
- Lee, D, & Swales, J. (2006). A corpus based EAP course for NNS doctoral students: Moving from Available specialized corpora to self-compiled corpora. *English for Specific purposes* 25, 56-75.
- Leech, G. (1992). Corpora and theories of linguistic performance. In J. Svartvik (Ed.) *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82* (pp. 105-122). Berlin: Mouton de Gruyter.
- . (1997). Teaching and language corpora: a convergence. In: A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles (Eds.), *Teaching and language Corpora* (pp. 1-23). London: Longman.
- Le Foll, E. (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* (3rd edition). <https://elenlefol.pressbooks.com> (Accessed 13 March 2024).
- Lexical Computing. (2017). *Sketch engine* [Computer software] <https://www.sketchengine.eu/>.
- . (2023). *SKELL - Sketch engine for language learning* [Computer software]. <https://skell.sketchengine.eu/#home?lang=en>.
- Leńko-Szymańska, A. (2008). Non-native or non-expert? The use of connectors in native and foreign learners' texts. *Acquisition et interaction en langue étrangère*, 27, 91-108.

- Leńko-Szymańska, A. (2015). The English vocabulary profile as a benchmark for assigning levels to learner corpus data. In M. Callies & S. Götz (Eds.), *Learner Corpora in Language Testing and Assessment* (pp. 115-140). Amsterdam: John Benjamins.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Lewis, Michael. (2002). *The Lexical Approach - the state of ELT and a way forward*. Boston: Thomson Heinle.
- Lippi-Green, R. (2012). *English with an accent* (2nd edition). New York: Routledge.
- Liu, D. (2012). The most frequently-used multi-word constructions in academic written English: A multi-corpus study. *English for Specific Purposes*, 31, 25-35.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of Second Language Acquisition* (pp. 413-468). Cambridge: Academic Press.
- Lopriore, L., & Vettorel, P. (2015). Promoting awareness of Englishes and ELF in the English language classroom. In H. Bowles & A. Cogo (Eds.), *International Perspectives on English as a Lingua Franca: Pedagogical Insights* (pp. 13-34). London: Palgrave Macmillan.
- Lorenz, G. (1999). *Adjective Intensification – Learners versus Native Speakers - A corpus study of argumentative writing*. Amsterdam: Rodopi.
- Linnarud, M. (1975). Lexis in free production: An analysis of the lexical texture of Swedish students' written work. *Swedish-English Contrastive Studies*, 6.
- Lüsebrink, H.-J. (2005). *Interkulturelle Kommunikation: Interaktion, Fremdwahrnehmung, Kulturtransfer*. Stuttgart: J. B. Metzler.
- . (2016). *Interkulturelle Kommunikation - Interaktion, Fremdwahrnehmung, Kulturtransfer* (4th edition). Stuttgart: J.B. Metzler.
- Ma, Q., Chiu, M. M., Lin, S., & Mendoza, N. B. (2023). Teachers' perceived corpus literacy and their intention to integrate corpora into classroom teaching: A survey study. *ReCALL*, 35(1), 19-39. <https://doi.org/10.1017/S0958344022000180>
- Ma, Q., Tang, J., & Lin, S. (2021). The development of corpus-based language pedagogy for TESOL teachers: A two-step training approach facilitated by online collaboration. *Computer Assisted Language Learning*, 35(4), 2731-2760. <https://doi.org/10.1080/09588221.2021.1895225>.
- Maffesoli, M. (1996). *Ordinary knowledge: Introduction to interpretive sociology*. Cambridge: Polity Press.
- Mauranen, A. (2006). Signaling and preventing misunderstanding in English as lingua franca communication. *International Journal of the Sociology of Language*, 177, 123-150.
- . (2012). *Exploring ELF: Academic English shaped by non-native speakers*. Cambridge: Cambridge University Press.
- Mauranen, A., Carey, R., & Ranta, E. (2015). New answers to familiar questions: English as a lingua franca. In D. Biber, & R. Reppen (Eds.), *The Cambridge Handbook of English Corpus Linguistics* (pp. 401-417). Cambridge: Cambridge University Press.
- McCarthy, M. (1991). *Discourse analysis for language teachers*. Cambridge: Cambridge University Press.
- . (2001). *Issues in applied linguistics*. Cambridge: Cambridge University Press.
- McCarthy, M., Matthiessen, C., & Slade, D. (2013). Discourse analysis. In N. Schmitt (Ed.), *An introduction to applied linguistics* (2nd edition, pp. 53-69). London: Routledge.
- McEney, T., & Hardie, A. (2012). *Corpus linguistics*. Edinburgh: Edinburgh University Press. <http://corpora.lancs.ac.uk/clmtp/index.php> (Accessed 10 March 2017).
- McEney, A., & Xiao, R. (2004). *Developing linguistic corpora: a guide to good practice*. Lancaster University. <https://users.ox.ac.uk/~martinw/dlc/chapter4.htm#section1> (Accessed 12 February 2024).
- Medienpädagogischer Forschungsverbund Südwest (Ed.). (2023). *JIM-Studie 2023*. Stuttgart. https://www.mpfs.de/fileadmin/files/Studien/JIM/2022/JIM_2023_web_final_kor.pdf (Accessed 12 March 2024).

- Meierkord, C. (2012). *Interactions across Englishes: Linguistic Choices in Local and International Contact Situations*. Cambridge: Cambridge University Press.
- . (2000). Interpreting successful lingua franca interaction. An analysis of non-native/non-native small talk conversations in English. *Linguistik online*, 5(1). <https://bop.unibe.ch/linguistik-online/article/download/1013/1673?inline=1> (Accessed 03 May 2024).
- Meunier, F., & Granger, S. (Eds.). (2008). *Phraseology in foreign language learning and teaching*. Amsterdam: John Benjamins.
- Meyer, M., Volkman L., & Grimm, N. (2022). *Teaching English* (2nd edition). Tübingen: Narr Francke Attempto.
- Meyer, R. (2004). *Lehren Kompakt*. Stuttgart: hep. https://sts-ghrf-giessen.bildung.hessen.de/intern/Schriftliche_Vorbereitung/Grundschule/Verben_zur_Formulierung_von_Lernzielen.pdf (Accessed 15 March 2023).
- MICASE. (2002-2019). *Michigan Corpus of Academic Spoken English* [Corpus]. Rita Simpson-Vlach, R.; Swales, J. Briggs, S., Römer, U. University of Michigan - English Language Institute. <https://quod.lib.umich.edu/cgi/c/corpus/corpus?c=micase;page=simple>.
- Mihajlović, D. (2017). *Kommunikation, Kollaboration, Kreativität und kritisches Denken – mehr als Buzzwords*. [Blog] <https://mihajlovicfreiburg.com/2017/04/18/kommunikation-kollaboration-kreativitaet-und-kritisches-denken-mehr-als-buzzwords/> (Accessed 10 March 2023).
- Milton, J., & Tsang, E. S.C. (1993). A corpus-based study of logical connectors in EFL students' writing: Directions for future research. In R. Pemberton & E. S. C. Tsang (Eds.), *Lexis in studies* (pp. 215-246). Hong Kong: Hong Kong University Press.
- Mindt, D. (1996). English corpus linguistics and the foreign language teaching syllabus. In J. Thomas & M. Short (Eds.), *Using Corpora for language research* (pp. 232–247). Harlow: Longman.
- Ministerium für Bildung des Landes Sachsen-Anhalt. (2022). *Fachlehrplan Gymnasium Englisch*. https://lisa.sachsen-anhalt.de/fileadmin/Bibliothek/Politik_und_Verwaltung/MK/LISA/Unterricht/Lehrplaene/Gym/Anpassung/Englisch_FLP_Gym_01_07_2019.pdf (Accessed 12 December 2023).
- Ministerium für Bildung, Jugend und Sport Brandenburg. (2022). *Rahmenlehrplan für die gymnasiale Oberstufe Teil C Englisch*. https://bildungsserver.berlin-brandenburg.de/fileadmin/bbb/unterricht/rahmenlehrplaene/gymnasiale_oberstufe/curricula/2022/Teil_C_RLP_GOST_2022_Englisch.pdf (Accessed 12 December 2023).
- Ministerium für Bildung und Kultur Saarland. (2015). *Lehrplan Englisch Gymnasiale Oberstufe Hauptphase Grundkurs*. https://www.saarland.de/SharedDocs/Downloads/DE/mbk/Lehrplaene/Lehrplaene_GOS_ab_2019_2020/Englisch/LP_En_GOS_HP_GK_2015-2019.pdf?__blob=publicationFile&v=3 (Accessed 12 December 2023).
- . (2023). *Englisch Fortgeführte Fremdsprache Lehrplan Gymnasiale Oberstufe Hauptphase Leistungskurs und Grundkurs*. https://www.saarland.de/SharedDocs/Downloads/DE/mbk/Lehrplaene/Lehrplaene_GOS_ab_2019_2020/Englisch/LP_EN_HP_GOS_2023.pdf?__blob=publicationFile&v=2 (Accessed 12 December 2023).
- Ministerium für Bildung, Wissenschaft und Kultur Mecklenburg-Vorpommern. (2019). *Rahmenlehrplan für die Qualifikationsphase der gymnasialen Oberstufe Englisch*. https://www.bildung-mv.de/export/sites/bildungsserver/downloads/unterricht/rahmenplaene_allgemeinbildende_schulen/Englisch/RP_EN_SEK2.pdf (Accessed 12 December 2023).
- Ministerium für Bildung, Wissenschaft und Weiterbildung Rheinland-Pfalz. (1998). *Lehrplan Englisch*. <https://lehrplaene.bildung-rp.de> (Accessed 12 December 2023).
- Ministerium für Bildung, Wissenschaft, Weiterbildung und Kultur Rheinland Pfalz. (2014). *Lehrplan Englisch*. <https://lehrplaene.bildung-rp.de> (Accessed 12 December 2023).
- Ministerium für Bildung, Wissenschaft und Kultur des Landes Schleswig-Holstein. (2014). *Fachanforderungen Englisch Allgemein bildende Schulen Sekundarstufe I / Sekundarstufe II*.

- <https://fachportal.lernnetz.de/sh/fachanforderungen/englisch.html> (Accessed 12 December 2023).
- Ministerium für Kultus, Jugend und Sport Baden-Württemberg. (2016). *Bildungsplan des Gymnasiums - Bildungsplan Englisch als erste Fremdsprache*. <https://www.bildungsplaene-bw.de/Lde/LS/BP2016BW/ALLG/GYM/E1> (Accessed 12 December 2023).
- Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen. (2023). *Kernlehrplan für die Sekundarstufe II Gymnasium/Gesamtschule in Nordrhein-Westfalen - Englisch*. https://www.schulentwicklung.nrw.de/lehrplaene/lehrplan/329/gost_klp_e_2023_06_07.pdf (Accessed 12 December 2023).
- Mondada, L. (2014). Pointing, talk, and the bodies. In M. Seyfeddinipur & M. Gullberg (Eds.), *From gesture in conversation to visible action as utterance: Essays in honor of Adam Kendon* (pp. 95-124). Amsterdam: John Benjamins.
- Moore, D. (2003). *Les représentations des langues et de leur apprentissage*. Paris: Didier.
- Moran, P. (2001). *Teaching culture: perspectives in practice*. Boston: Heinle & Heinle.
- Mori, J. (2002). Task design, plan, and development of talk in interaction: An analysis of a small group activity in a Japanese language classroom. *Applied Linguistics*, 23(3), 323–47.
- Morris, L., & Cobb, T. (2004). Vocabulary profiles as predictors of the academic performance of Teaching English as a Second Language trainees. *System*, 32(1), 75-87.
- Morrow, K. (1977). Authentic texts and ESP. In S. Holden (Ed.), *English for Specific Purposes* (pp. 13-17). London: Modern English Publications.
- Moscovici, S. (1961). *La psychoanalyse, son image, son public*. Paris: Presses Universitaires de France.
- Mukherjee, J. (2002). *Korpuslinguistik und Englischunterricht: eine Einführung*. Frankfurt am Main: Peter Lang.
- . (2004). Bridging the gap between applied corpus linguistics and the reality of English language teaching in Germany. In U. Connor & T. A. Upton (Eds.), *Applied Corpus Linguistics. A Multi-dimensional perspective* (pp. 239-250). Rodopi: Amsterdam.
- . (2006). Corpus linguistics and language pedagogy: The state of the art - and beyond. In S. Braun, J. Kohn & J. Mukherjee, J. (Eds.), *Corpus technology and language pedagogy* (pp. 5-24). Frankfurt am Main: Peter Lang.
- Muß-Merholz, J. (2017). *Die 4K-Skills: Was meint Kreativität, kritisches Denken, Kollaboration, Kommunikation?* <https://www.joeran.de/die-4k-skills-was-meint-kreativitaet-kritisches-denken-kollaboration-kommunikation/> (Accessed 10 March 2023).
- Myers, G. (2010). Stance-taking and public discussion in blogs. *Critical Discourse Studies*, 7 (4), 263-275.
- Nasseri, M., & Thompson, P. (2021). Lexical density and diversity in dissertation abstracts: Revisiting English L1 vs. L2 text differences. *Assessing Writing*, 47, 100511.
- Niedersächsisches Kultusministerium. (2017). Kerncurriculum für das Gymnasium - gymnasiale Oberstufe - Englisch. https://cuvo.nibis.de/cuvo.php?p=search&k0_0=Schulbereich&v0_0=Sek+II&k0_1=Schulform&v0_1=Gymnasiale+Oberstufe&k0_2=Fach&v0_2=Englisch& (Accessed 12 December 2023).
- Norrick, N. (2013). *Transcription conventions*. Department of English Linguistics, Saarland University. http://www.uni-saarland.de/fileadmin/user_upload/Professoren/fr43_Engling/downloads/trans_conv.pdf (Accessed 17 May 2014).
- O'Dowd, R., & Lewis, T. (Eds.). (2016). *Online intercultural exchange: Policy, pedagogy, practice*. London: Routledge.
- O'Halloran, K. L. (2011). Multimodal Discourse Analysis. In K. Hyland & B. Paltridge (Eds.), *Bloomsbury Companion to Discourse Analysis* (pp. 120-137). London: Bloomsbury.
- O'Keefe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom*. Cambridge: Cambridge University Press.
- O'Sullivan, I. (2007). Enhancing a process-oriented approach to literacy and language learning: The role of corpus consultation literacy. *ReCALL*, 19(3), 269–286.
- Olson, G. M. & Olson, J. S. (2000). Distance matters. *Human-Computer Interaction*, 15(2-3), 139-178.

- Paikeday, T.M. (1985). *The Native Speaker is Dead!* Toronto: Paikeday.
- Paltridge, B. (2012). *Discourse analysis: An introduction* (2nd edition). London: Bloomsbury Publishing.
- Paradis, C. (2003). Between epistemic modality and degree: The case of really. In R. Facchinetti, F. Palmer & M. Krug (Eds.), *Modality in Contemporary English* (pp. 191-222). Berlin, Boston: De Gruyter Mouton.
- Pérez-Paredes, P. (2019). A systematic review of the uses and spread of corpora and data-driven learning in CALL research during 2011–2015. *Computer Assisted Language Learning*, 35(1-2), 36-61.
- Pérez-Paredes, P., & Bueno-Alastuey, M. C. (2019). A corpus-driven analysis of certainty stance adverbs: Obviously, really and actually in spoken native and learner English. *Journal of Pragmatics*, 140, 22-32.
- Pfiffner, M., Sterel, S., & Hassler, D. (2021). *4K und digitale Kompetenzen – Chancen und Herausforderungen* (Volume 1). Bern: hep-Verlag.
- Phipps, A.M., & Gonzalez, M. (2004). *Modern Languages: Learning and Teaching in an Intercultural Field*. London: Sage.
- Piątkowska, K. (2015). From cultural knowledge to intercultural communicative competence: Changing perspectives on the role of culture in foreign language teaching. *Intercultural Education*, 26(5): 397–408.
- Pietilä, P. (2015). Lexical diversity in L2 academic writing: a look at MA thesis conclusions. In P. Pietilä, K. Doró & R. Pípalová (Eds.), *Lexical issues in L2 writing* (pp. 105-126). Newcastle upon Tyne. Cambridge Scholars Publishing.
- Pietikäinen, K. S. (2017). *English as a Lingua Franca in Intercultural Relationships: Interaction, Identity, and Multilingual Practices of ELF Couples* (Unpublished doctoral dissertation). University of Helsinki.
- Pike, K. (1967). *Language in relation to a unified theory of the structure of human behaviour*. The Hague: Mouton.
- Piller, I. (2001). Who, if anyone, is a native speaker? *Anglistik*, 12(2), 109-121.
- Pinto, T.P., Crosthwaite, P., de Carvalho, C. T., Spinelli, F., Serpa, T., Garcia, W., & Ottaiano, A. O. (Eds.) (2023). *Using language data to learn about language: a teachers' guide to classroom corpus use*. Brisbane: The University of Queensland. <https://doi.org/10.14264/3bbe92d> (Accessed 25 February 2024)
- Pitzl, M.-L. (2010). *ELF in International business: Resolving miscommunication and reaching shared understanding*. Saarbrücken: VDM.
- Pitzl, M.-L. (2016). Expanding frontiers: Prospects on the creativity of ELF. In M.-L. Pitzl, & R. Osimk-Teasdale (Eds.), *English as a lingua franca: Perspectives and prospects. Contributions in honour of Barbara Seidlhofer* (pp. 275–279). Berlin, Boston: De Gruyter Mouton.
- Poitzmann, N., & Sobel, M. (2024). *Upgrade: 21st Century Skills: Das 4K-Modell des Lernens in der Praxis*. Hannover: Friedrich Verlag.
- Poole, R. (2020). “Corpus can be tricky”: revisiting teacher attitudes towards corpus-aided language learning and teaching. *Computer Assisted Language Learning*, 35(7), 1620–1641. <https://doi.org/10.1080/09588221.2020.1825095>.
- Poyatos, F. (1992). Non-verbal communication in foreign language teaching: Theoretical and methodological perspectives. In A. Helbo (Ed.), *Evaluation and Language Teaching* (pp. 115-143). Bern: Peter Lang.
- Prodromou, L. (2005). “You see, it’s sort of tricky for the L2-user”: *The puzzle of idiomaticity in English as a lingua franca* (Unpublished doctoral dissertation). University of Nottingham.
- Puenteadura, R. R. (2006). *Transformation, technology, and education*. [Power Point Presentation] http://hippasus.com/resources/tte/puenteadura_tte.pdf (4 January 2024).
- Queen, R. (2015). *Vox popular: The surprising life of language in the media*. West Sussex: John Wiley & Sons.
- Quinn, B. (2013). Arts festivals, urban tourism, and cultural policy. *Journal of Policy Research in Tourism, Leisure and Events*, 5(3), 236-254. <https://doi.org/10.1080/19407963.2013.776372>.

- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. New York: Longman.
- Raymond, G. (2004). Prompting action: The stand-alone “so” in ordinary conversation. *Research on Language and Social Interaction*, 37, 185-281.
- Redeker, G. (2006). Discourse markers as attentional cues at discourse transitions. In K. Fischer (Ed.), *Approaches to Discourse Particles* (pp. 339-358). Amsterdam: Elsevier.
- Renouf, A. (1997). Teaching corpus linguistics to teachers of English. In A. Wichmann, T. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Applied corpus linguistics. A multidimensional perspective* (pp. 251-265). London: Rodopi.
- Reppen, R. (2010). *Using corpora in the language classroom*. Cambridge: Cambridge University Press.
- Research and Development Unit for English Studies, Birmingham City University. (1999-2024). *WebCorpLSE* [Computer software] <https://www.webcorp.org.uk/wcx/lse/corpora>
- Richards, J. C. (2006). *Communicative language teaching today*. Cambridge: Cambridge University Press.
- Risager, K. (2006). *Language and culture: Global flows and local complexity*. Clevedon: Multilingual Matters.
- Roberts, C., Byram, M., Barro, A., Jordan, S., & Street, B. (2001). *Language learners as ethnographers: Introducing cultural process into advanced language learning*. Clevedon: Multilingual Matters.
- Römer, U. (2005). *Progressives, patterns, pedagogy. A corpus-driven approach to English progressive forms, functions, contexts and didactics*. Amsterdam: John Benjamins.
- . (2011). Corpus research applications in second language teaching. *Annual Review of Applied Linguistics*, 31, 205-225.
- Ros i Solé, C. (2013). Cosmopolitan speakers and their cultural cartographies. *The Language Learning Journal*, 41(3), 326-39.
- Rüdiger, S., & Mühleisen, S. (Eds.). (2020). *Talking about Food: The social and the global in eating communities* (Volume 47). Amsterdam: John Benjamins.
- Rutaganda, E. (2024). The role of cultural festivals in promoting social cohesion and cultural understanding. *International Journal of Humanity and Social Sciences*, 3, 13-25.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, 696-735.
- Sächsisches Staatsministerium für Kultus. (2022). *Lehrplan Gymnasium Englisch*. https://www.schulportal.sachsen.de/lplandb/index.php?lplanid=816&lplansc=Ot016RpUeeLWYkna0Y8m&token=a37bd7635bd3f198cd23fd730fe2d3b6#page816_12348 (Accessed 12 December 2023).
- Samovar, L. A. & Porter, R. E. (Eds.). (1993). *Intercultural communication: a reader*. Belmont: Wadsworth.
- Santino, J. (1996). *New old-fashioned ways - holidays and popular culture*. Knoxville: University of Tennessee Press.
- Savignon, S. J. (2018). Communicative competence. *The TESOL encyclopedia of English language teaching*, 1-7. DOI:10.1002/9781118784235.eelt0047.
- Schegloff, E. (1982). Discourse as an interactional achievement: Some uses of uh huh and other things that come between sentences. In D. Tannen (Ed.), *Analyzing discourse: Text and talk* (pp. 71-93). Washington, DC: Georgetown University Press.
- . (2007). *Sequence organization in interaction: A primer in conversation Analysis I*. Cambridge: Cambridge University Press.
- Schegloff, E., & Sacks, H. (1973). Opening Up Closings. *Semiotica*, 8, 289-327.
- Schiffrin, D. (1987). *Discourse markers*. Cambridge, U.K.: Cambridge University Press.
- Schleicher, A. (2013, May 8). *21st Century Skills* [Keynote re:publica 2013. Youtube. https://www.youtube.com/watch?v=lbb5KE6Cl_w (Accessed 10 March 2023).
- Schmidt, S. (2020). *Rapport management in online spoken interaction: A cross-cultural linguistic analysis of communicative strategies* [Doctoral dissertation]. Birmingham City University. https://www.open-access.bcu.ac.uk/12132/1/Thesis%20Selina%202020_07_08.pdf (Accessed 12 February 2023).

- Schneider, E. W. (2011): *English around the world*. An Introduction. Cambridge: Cambridge University Press.
- Schugk, M. (2004). *Interkulturelle Kommunikation. Kulturbedingte Unterschiede in Verkauf und Werbung*. München: Verlag Franz Vahlen.
- Scollon, R., & LeVine, P. (2004). Multimodal Discourse Analysis as the Confluence of Discourse and Technology. In P. LeVine & R. Scollon (Eds.), *Discourse and technology: Multimodal discourse analysis* (pp. 1-6). Washington: Georgetown University Press.
- Seedhouse, P. (2005). Conversation analysis and language learning. *Language Teaching*, 38(4), 165-187.
- Scott, M. (1999). *WordSmith* [Computer software]. <https://www.lexically.net/wordsmith/>.
- Seidlhofer, B. (2003). *A concept of international English and related issues: From 'real English' to 'realistic English'*. Strasbourg: Council of Europe.
- Seidlhofer, B. (2011). *Understanding English as a lingua franca*. Oxford: Oxford University Press.
- Seidlhofer, B., & Widdowson, H. (2017). Competence, capability and virtual language. *Lingue e Linguaggi*, 24, 23-36.
- Senatorin für Bildung und Wissenschaft. (2015). *Bildungsplan Fortgeführte moderne Fremdsprachen*. <https://www.lis.bremen.de/schulqualitaet/curriculumentwicklung/bildungsplaene/sekundarbereich-ii-allgemeinbildend-16698#> (Accessed 12 December 2023).
- Senatsverwaltung für Bildung, Jugend und Familie Berlin. (2022). *Rahmenlehrplan für die gymnasiale Oberstufe Teil C Englisch*. <https://www.berlin.de/sen/bildung/unterricht/faecher-rahmenlehrplaene/rahmenlehrplaene/> (Accessed 12 December 2023).
- Shabani, K., & Zanussi, M. P. (2015). The impact of watching captioned TV series on vocabulary development of EFL students. *Journal for the Study of English Linguistics*, 3(1), 118-129.
- Shakouri, E., & Shakouri, N. (2014). On the death of native speaker: A revisiting stance. *International Journal of Educational Investigations*, 1(1), 220-226.
- Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, 31(4), 487-512.
- Sinclair, J. M. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- . (1995). *Collins COBUILD English dictionary*. Glasgow: HarperCollins.
- . (1997). Corpus evidence in language description. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 27-39). London: Longman.
- . (2003). *Reading concordances*. London: Longman.
- . (2005). Corpus and text - basic principles. In M. Wynne (Ed.), *Developing linguistic corpora: a guide to good practice* (pp. 1-16). Oxford: Oxbow Books.
- Sinclair, J. M., & Renouf, A. (1988). A lexical syllabus for language learning. In R. Carter & M. McCarthy (Eds.), *Vocabulary in language teaching* (pp. 140-158). London: Longman.
- Sindoni, M. G. (2023). *Multimodality and translanguaging in video interactions*. Cambridge: Cambridge University Press.
- Spencer-Oatey, H. (2002). Managing rapport in talk: Using rapport sensitive Incidents to explore the motivational concerns underlying the management of relations. *Journal of Pragmatics*, 34(5), 529-545.
- Spencer-Oatey, H., & Harsch, C. (2016). The critical incident technique. In Z. Hua (Ed.), *Research methods in intercultural communication: a practical guide* (p. 223-238). Oxford: Wiley-Blackwell Publishing.
- Staatsinstitut für Schulqualität und Bildungsforschung München. (2023). *Lehrplan Englisch 12/13*. <https://www.lehrplanplus.bayern.de/fachlehrplan/gymnasium/12/englisch/grundlegend> (Accessed 12 December 2023).
- Stadler, F. (2016). *Kultur der Digitalität*. Berlin: Suhrkamp.
- Stefanowitsch, A. (2019). *Corpus linguistics: A guide to the methodology*. Berlin: Language Science Press.
- Stenström, A. B. (1986). What does really really do? In G. Tottie & I. Bäcklund (Eds.), *English in speech and writing: A symposium* (pp. 149-163). Stockholm: Almqvist & Wiksell.
- Storey, J. (2006). *Cultural theory and popular culture - an introduction*. Harlow: Pearson.

- Surkamp, C., & Viebrock, B. (Eds.).(2018). *Teaching English as a foreign language*. Stuttgart: J.B. Metzler.
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16(3), 371-391.
- Swales, J., & Malczewski, B. (2001). Discourse management and new-episode flags in MICASE. In R. Simpson & J. Swales (Eds.), *Corpus Linguistics in North America: Selections from the 1999 Symposium* (pp. 145-164). Ann Arbor: University of Michigan Press.
- Swann, W., & Bosson, J. (2008). Identity negotiation: A theory of self and social Interaction. In O. P. John, R. W. Robins & L. Pervin (Eds.), *Handbook of Personality: Theory and Research* (3rd edition, pp. 448–471). New York: Guilford Press.
- TaCoCASE. (2023). *Transatlantic Component of the CASE project* [Corpus]. Collet, C. Saarland University & Trier University of Applied Sciences. <http://umwelt-campus.de/case/tacocase> (Accessed 13 March 2024).
- Tang, J. C., & Isaacs, E. (1993). Why do users like video? Studies of multimedia-supported collaboration. *Computer Supported Cooperative Work* 1, 163–196.
- Tarone, E. (1981). Some thoughts on the notion of communication strategy. *TESOL Quarterly*, 15(3), 285–295.
- Teubert, W. (2005). My version of corpus linguistics. *International Journal of Corpus Linguistics*, 10 (1), 1-13.
- The CASE project. (2017a). *CASE transcription conventions*. Birkenfeld: Trier University of Applied Sciences.[https://www.umwelt-campus.de/fileadmin/Umwelt-Campus/SK-Weiterbildung/Dateien/CASE Transcription conventions 0817.pdf](https://www.umwelt-campus.de/fileadmin/Umwelt-Campus/SK-Weiterbildung/Dateien/CASE_Transcription_conventions_0817.pdf). (Accessed 11 November 2017).
- . (2017b). *CASE Taxonomy for non-verbal elements*. Birkenfeld: Trier University of Applied Sciences. <http://umwelt-campus.de/case-conventions> (Accessed 11 November 2017).
- Thomas, A. (1991). Kulturstandards in der internationalen Begegnung. *Sozialwissenschaftlicher Studienkreis für internationale Probleme* 61. Saarbrücken/ Fort Lauderdale: Breitenbach.
- . (1993): Psychologie interkulturellen Lernens und Handelns. In A. Thomas (Ed.), *Kulturvergleichende Psychologie* (pp. 377-424). Göttingen: Hogrefe.
- Thompson, P. (2005). Spoken language corpora. In M. Wynne (Ed.), *Developing Linguistic Corpora: A Guide to Good Practice*. <http://users.ox.ac.uk/~martinw/dlc/chapter5.htm> (Accessed 17 August 2022).
- Thornbury, S. (2002). *How to teach vocabulary*. Edinburgh: Pearson Education Limited.
- Thornbury, S., & Slade, D. (2006). *Conversation: From Description to Pedagogy*. Cambridge: Cambridge University Press.
- Thüringer Ministerium für Bildung, Jugend und Sport (2019). *Lehrplan für den Erwerb der allgemeinen Hochschulreife - Englisch*. <https://www.schulportal-thueringen.de/media/detail?tspi=1395> (Accessed 12 December 2023).
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam: John Benjamins.
- . (2010). Theoretical overview of the evolution of corpus linguistics. In A. O’Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpora and English Language Teaching and Learning* (pp. 14-28). New York: Routledge.
- Tolins, J., & Tree, J. E. F. (2014). Addressee backchannels steer narrative development. *Journal of Pragmatics*, 70, 152-164.
- Tsui, A. B. (1991). The pragmatic functions of *I don’t know*. *Text*, 11, 607–622.
- UCREL. (2018a). *CLAWS Part-of-Speech Tagger for English* [Computer software] <http://ucrel.lancs.ac.uk/claws/>.
- UCREL. (2018b). *Semantic Analysis System (USAS)* [Computer software] <https://ucrel.lancs.ac.uk/usas/>.
- Ure, J. (1971). Lexical density and register differentiation. In G. E. Perren & J. L. M. Trim (Eds.), *Applications of linguistics. Selected papers of the Second International Congress of Applied Linguistics, Cambridge 1969* (pp. 443-452). Cambridge: Cambridge University Press.

- Van Ek, J. A. (1986). Objectives for foreign language learning (Volume 1). Strasbourg: Council of Europe.
- Vatter, C., & Zapf, E. (2012). *Interkulturelle Kompetenz: Erkennen – verstehen – handeln. Französisch*. Stuttgart: Ernst Klett.
- Vettorel, P. (2015). World Englishes and English as a lingua franca: Implications for teacher education and ELT. *Iperstoria*, 6, 229- 244.
- ViMELF. (2018). *Corpus of video-mediated English as lingua franca conversations* [Corpus]. The CASE project. Trier University of Applied Sciences. <http://umwelt-campus.de/case> (Accessed 13 March 2024).
- Vo, S. (2019). Use of lexical features in non-native academic writing. *Journal of Second Language Writing*, 44, 1-12.
- Vodafone Stiftung Deutschland (Ed.). (2023). *Zwischen Hoffnung und Sorge - Digitale Bildung an Schulen in Deutschland und Europa aus Sicht von Eltern*. <https://www.vodafone-stiftung.de/ipsos-elternbefragung-2023/> (Accessed 13 March 2023).
- VOICE. (2013). *The Vienna-Oxford International Corpus of English* (version 2.0) [Corpus]. Seidlhofer, B. University of Vienna. <https://www.univie.ac.at/voice/>.
- Wagner, J., & Firth, A. (1997). Communication strategies at work. In G. Kasper, & E. Kellerman (Eds.), *Communication strategies: Psycholinguistic and sociolinguistic perspectives* (pp. 323-344). London: Routledge.
- Walther, J. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication research*, 23, 3-43.
- Walther, J. B. (2011). Theories of computer-mediated communication and interpersonal relations. *The handbook of interpersonal communication*, 4, 443-479.
- Warschauer, M. & Grimes, D. (2007). Audience, authorship, and artifact: The emergent semiotics of web 2.0. *Annual Review of Applied Linguistics*, 27, 1-23.
- Weinert, F. E. (Ed.) (2001). *Leistungsmessung in Schulen*. Weinheim und Basel: Beltz.
- Welsch, W. (1992). Rede zur Kultur. In F. Köllmayr (Ed.), *Soblau* (pp. 16-30). München: A1 Verlag.
- . (2010). Was ist eigentlich Transkulturalität? In L. Darowska & C. Machold (Eds.), *Hochschule als transkultureller Raum? Beiträge zu Kultur, Bildung und Differenz* (pp. 39–66). Bielefeld: Transkript.
- Werner, V. (2018). Linguistics and pop culture: Setting the scene(s). In V. Werner (Ed.), *The language of pop culture* (pp. 3-26). New York: Routledge.
- . (2022). Pop culture in applied linguistics: International perspectives. *International Journal of Applied Linguistics*, 173(2), 161-171.
- Werner, V., & Tegge, F. (Eds.). (2020). *Pop culture in language education: Theory, research, practice* (1st edition). New York: Routledge. <https://doi.org/10.4324/9780367808334>.
- Widdowson, H. G. (2012). ELF and the inconvenience of established concepts. *Journal of English as a Lingua Franca*, 1(1), 5–26.
- Wilkinson, J. (2020). From native speaker to intercultural speaker and beyond. Jackson, J. (ed). *The Routledge Handbook of Language and Intercultural Communication* (2nd edition, pp. 283-298). New York: Routledge.
- Williams, R. 1981. *Culture*. London: Fontana.
- . (1983). *Keywords*. London: Fontana.
- Wray, A., Trott, K., & Bloomer, A. (1998). *Projects in Linguistics: A practical Guide to researching language*. London: Arnold.
- Wynne, M. (Ed.). (2006). *Developing linguistic corpora—a guide to good practice*. EADH: The European Association for Digital Humanities. https://icar.cnrs.fr//ecole_thematique/contaci/documents/Baude/wynne.pdf (Accessed 10 June 2019).
- Yule, G. (1995). The paralinguistics of reference: Representation in reported discourse. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (pp. 185-196). Oxford: Oxford University Press.
- . (2006). *The Study of Language* (3rd edition). Cambridge: Cambridge University Press.

Zhao, J. (2006). The communicative functions of gestures in L2 speech. *The Arizona Working Papers in Second Language Acquisition and Teaching*, 13, 1–17.

APPENDIX

- Native speaker questionnaire
- Non-native speaker questionnaire
- CASE conversation guide
- Skype recorder instructions
- CASE Transcription conventions
- CASE Transcription guidelines for non-verbal elements
- Overview of transcription conventions

Baseline Data CASE (Native Speakers)

* Required

1. **ID ***

If you don't know your ID, please ask your instructor! (Example: SB01)

Part ONE: Personal Data

2. **First Name ***

3. **Last Name ***

4. **E-Mail ***

5. **Gender ***

Mark only one oval.

☐ male

☐ female

6. **Nationality ***

7. **Age ***

Mark only one oval.

- ☐ 15-20
☐ 21-25
☐ 26-30
☐ 31-35
☐ 36-40
☐ 41-45
☐ 46-50
☐ 51-55
☐ 56-60
☐ >60

8. **University ***

9. **Course of studies ***

Check all that apply.

- ☐ Bachelor
☐ Master
☐ Teacher Candidate
☐ Doctoral Candidate
☐ Other: _____

10. **Major ***

11. **Minor**

12. **Year of studies ***

(1 year = 2 semesters or 3 trimesters)

13. **Native Language ***

14. Do you speak other foreign languages? *

Mark only one oval.

- ☐ yes *Skip to question 15.*
- ☐ no *Stop filling out this form.*

Part TWO: Language background

15. How many foreign languages do you speak? *

Mark only one oval.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ more than 3

16. First foreign language? *

Please indicate in parentheses what your level of proficiency is for that language. 1= beginner, 2= intermediate, 3= advanced, 4= proficient.
Example: German intermediate --> German (2)

17. Second foreign language?

Please indicate in parentheses what your level of proficiency is for that language. 1= beginner, 2= intermediate, 3= advanced, 4= proficient.
Example: Spanish beginner --> Spanish (1)

18. Third foreign language?

Please indicate in parentheses what your level of proficiency is for that language. 1= beginner, 2= intermediate, 3= advanced, 4= proficient.

19. Fourth foreign language?

Please indicate in parentheses what your level of proficiency is for that language. 1= beginner, 2= intermediate, 3= advanced, 4= proficient.

20. **Have you lived in a country with an official language OTHER than English? ***

(3 months or more)

Mark only one oval.

- ☐ yes *Skip to question 21.*
- ☐ no *Stop filling out this form.*

Time abroad

21. **How long have you lived in a country with an official language other than English? ***

(in months)

22. **Name the country where you have lived ***

23. **Which language did you speak all day? ***

Mark only one oval.

- ☐ English (your native language)
- ☐ official language of that country
- ☐ Other:

Baseline Data CASE (Non-Native speakers)

* Required

1. ID *

If you don't know your ID, please ask your instructor! (Example: SB01)

Part ONE: Personal Data

2. First Name *

3. Last Name *

4. E-Mail *

5. Gender *

Mark only one oval.

☐

Male

☐

Female

6. Nationality *

7. **Age ***

Mark only one oval.

- ☐ 15-20
- ☐ 21-25
- ☐ 26-30
- ☐ 31-35
- ☐ 36-40
- ☐ 41-45
- ☐ 46-50
- ☐ 51-55
- ☐ 56-60
- ☐ >60

8. **Course of Studies ***

Mark only one oval.

- ☐ Bachelor
- ☐ Master (incl. Magister)
- ☐ Teacher Candidate
- ☐ Doctoral Candidate
- ☐ Other: _____

9. **Major ***

10. **Minor ***

11. **Year of Studies ***

(1 year = 2 semesters)

12. **University ***

13. **Native Language ***

14. **Other languages spoken (besides English) ***

15. **English is your.....foreign language ***

Mark only one oval.

- ☐ first
- ☐ second
- ☐ third
- ☐ fourth or more

PART TWO: English Language Learning

16. **Name the country where you studied English first! ***

Please indicate in parantheses if it's your native country or not. (NC = native country ; NNC = NOT my native country!)

17. **How long have you been learning English? ***

(in years)

18. **How/Where did you learn English? ***

Check all that apply.

- ☐ at school
- ☐ private tutor
- ☐ at University
- ☐ self-study
- ☐ Other: _____

19. **Have you attended any English classes/courses in an English speaking country? ***

Mark only one oval.

0 1 2 3 4

not at all ☐ ☐ ☐ ☐ ☐ 4 or more courses

20. **Have you lived in an English speaking country ***

(3 months or more)

Mark only one oval.

☐ yes *Skip to question 21.*

☐ no *Skip to question 23.*

Life in an English speaking country

21. **How long have you lived in an English speaking country?**

(in months)

22. **Name the country where you have lived!**

Part THREE: Use of English

a) at University

23. **How often do you usually use English in your studies/at University? ***

Mark only one oval.

☐ every day

☐ several days a week

☐ occasionally

☐ Other: _____

24. **Which part of your English communication partners (at University) are native speakers of English? ***

0= not at all , 1= 1-25% , 2=26-50% , 3= 51-75% , 4= 76-100%

Mark only one oval.

0

1

2

3

4

none of them

☐☐☐☐☐

all of them

30. **Who do you usually communicate with in English?**

Check all that apply.

- ☐ your boss
- ☐ colleagues
- ☐ foreign partners
- ☐ customers
- ☐ Other: _____

31. **Which topics do you usually discuss with your partners in English?**

Check all that apply.

- ☐ related to work
- ☐ specialized topics
- ☐ general topics
- ☐ Other: _____

Use of English for personal purposes

32. **How often do you use English for personal purposes? ***

Mark only one oval.

- ☐ every day
- ☐ several days a week
- ☐ occasionally

33. **How do you use English in your every day life? ***

Check all that apply.

- ☐ reading (e.g. literature, newspapers and magazines, Internet documentaries etc.)
- ☐ listening/ watching (e.g. music, radio, TV etc.)
- ☐ face-to-face communication (e.g. friends, lecturers, colleagues, etc.)
- ☐ social networks (e.g. facebook, twitter, skype etc.)
- ☐ Other: _____

34. **Who do you usually communicate with in English? ***

Check all that apply.

- ☐ family (native speakers)
- ☐ family (non-native speakers)
- ☐ friends (native speakers)
- ☐ friends (non-native speakers)
- ☐ Other: _____

35. **What part of your communication partners are native speakers? ***

0= not at all , 1= 1-25% , 2=26-50% , 3= 51-75% , 4= 76-100%

Mark only one oval.

	0	1	2	3	4	
none of them	☐	☐	☐	☐	☐	all of them

36. **What topics do you usually discuss? ***

Check all that apply.

- ☐ general
- ☐ academic
- ☐ specialized (e.g. about music, sports, cars etc.)
- ☐ Other: _____

37. **What variety of English do you use/ prefer? ***

Mark only one oval.

- ☐ British English
 - ☐ American English
 - ☐ European English
 - ☐ Other: _____
-

CASE Skype Conversation Guide

1 Requirements: You will hold two conversations with a CASE conversation partner on a topic specified in the seminar. The conversations should be at least 30 minutes long, but feel free to end the conversation at any point should you feel uncomfortable or should problems arise. You will need a computer that runs Skype, or access to one. This includes built-in or external speakers, a camera and a microphone.

2 All sessions need to be recorded (otherwise we will obtain no data). Participants will need to set up the technical infrastructure for that and perform the actual recording. Below you will find easy step-by-step instructions for installing (free) software on your computer. These are add-ons that can run with Skype and record both audio and video when you start the conversation. **Make sure the software is recording** when you start the Skype conversation (this should be the default option).

Privacy Notice: All personal information will be anonymized according to university guidelines during the conversion of the data as part of the finished corpus. Video data, in particular, will be filtered to ensure participants' anonymity. By participating in the project and signing the consent form you agree that the data can be used, in anonymized form, in the CASE corpus for linguistic research.

3 You will receive name and email address of a conversation partner at the partner university; you can contact them and set up the talk to take place in the designated week. They will also receive your email. Please treat their data confidentially.

4 The duration of the talks should be at least 30 minutes. The topics for the conversations are pre-selected. This means that you could prepare some questions that interest you about that topic, and be prepared to answer. The topics prompts should help you get the conversation started; you can of course talk about other things too. The conversation should develop naturally.

1. **"So, what are you studying"** – Course of studies and job prospects
2. **"And what do you do all day?"** – Life as a university student
3. **"University here and there"** – Different academic cultures
4. **"The role and future of English"** – Lingua Franca & global attitudes
5. **"Should learning be virtual?"** – Mobile learning, online courses: pros & cons
6. **"How do you celebrate?"** – Cultures and traditions
7. **"Let's talk about food"** Eating habits and preferences – what do you like to eat?
8. **"What's on? Popular Culture"** – Talk about your favorite TV show

5 The recording software will record the conversation in a suitable data format. In order to receive credit, you will have to hand in the recording to your tutor by the specified deadline. **Please name the file according to your assigned conversation code** (e.g. 01SB01BI01). Usually the transfer will take place via USB stick, since the files will be too large to be sent by email.

Skype recording instructions

Callnote (free program)

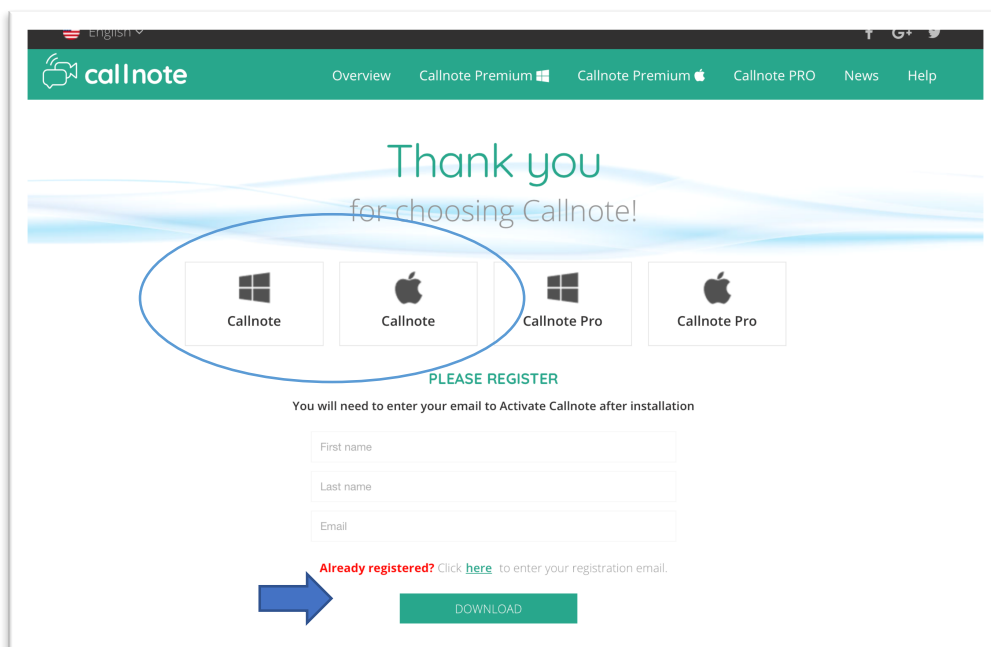
Download Link: <https://callnote.net/download-callnote-client/>

Follow the instructions on the homepage and **download "Callnote"** for free (not Callnote Pro as this has to be purchased).

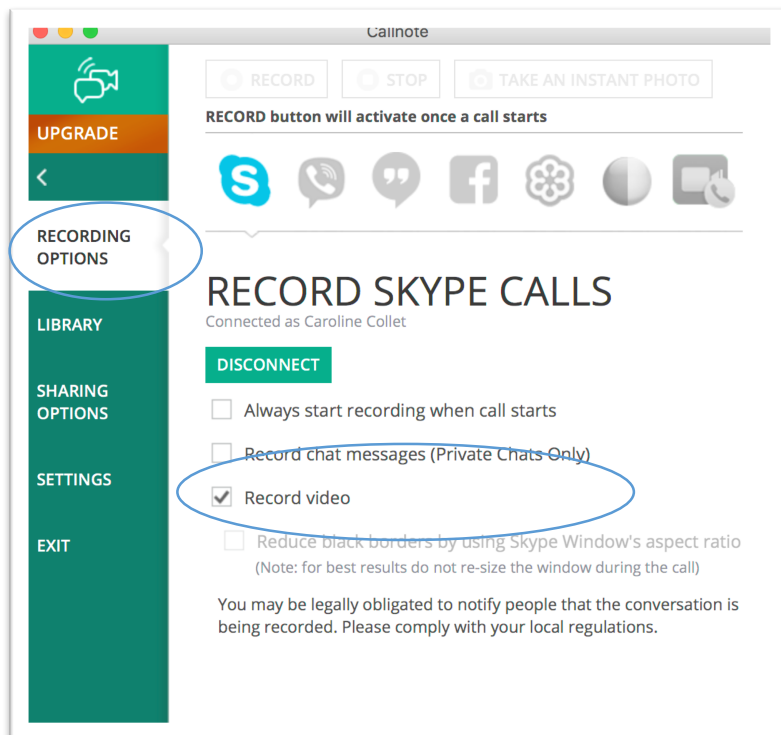
Log in on **Skype** and start the recording software **"Callnote"**. It will automatically connect to your Skype profile.

Before recording the actual conversation, we suggest to do a **test call / test recording** in order to see if Callnote is working fine. Make sure that Callnote is recording **both sides** of the conversation (we suggest to set "split screen") and also that you have **audio** and **video** output in your saved conversation file. Listen in to your saved file and make sure that the sound and video quality is high and everything is clear and comprehensible.

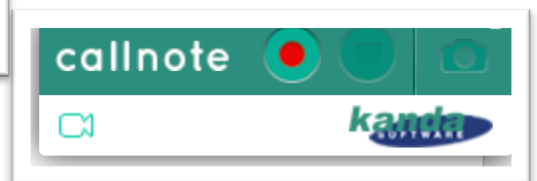
Screenshots for instructions from the website / program:



Recording your Skype conversations

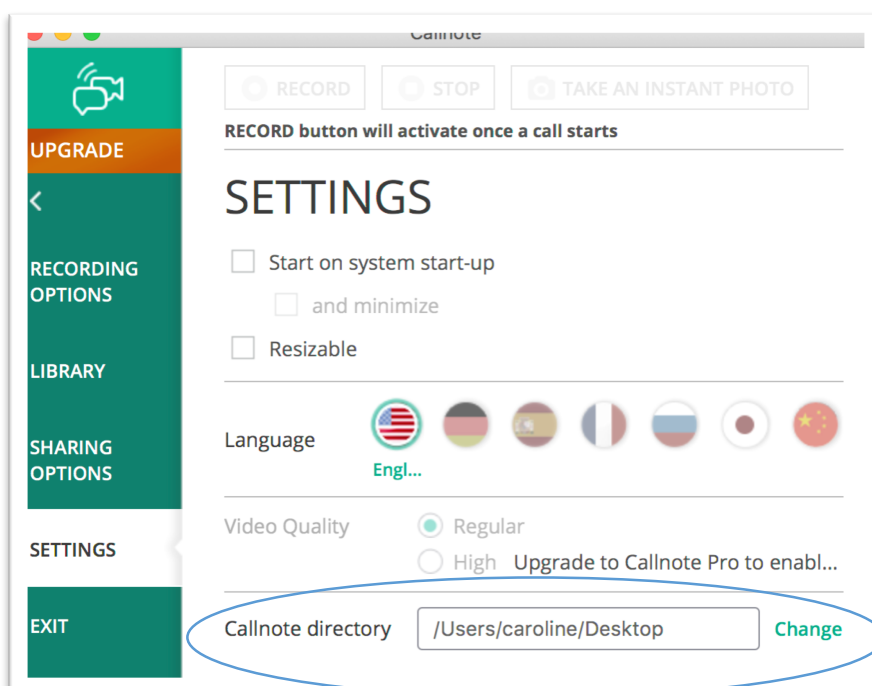


Start the recoring



Saving the output file

Settings >> Callnote directory >> select folder where you want to save the output file



CASE transcription conventions

General Transcription Conventions

Each line of transcription represents spoken language as segmented into intonation units. In English, an intonation unit typically consists of about five to seven words and expresses one new idea unit. Intonation units are likely to begin with a brief pause and to end in a clause-final intonation contour; they often match grammatical clauses. Each idea unit typically contains a subject, or given information, and a predicate, or new information; this flow from given to new information is characteristic of spoken language (Chafe 1994). Arranging each intonation unit on a separate line displays the frequency of intonation units beginning with *and* and thus highlights the greater fragmentation inherent in spoken language (Chafe 1982).

Capitalization is reduced to the pronoun *I* and proper names, diacritics are used to mark features of prosody rather than grammatical units, and non-lexical items, for example pause fillers like *ehm* and *um*, affirmative particles like *aha* or surprise markers like *oh*, are included. The specific transcription conventions are as follows.

she's out.	Period shows falling tone in the preceding element; suggesting finality.
oh yeah?	Question mark shows rising tone in the preceding element.
so, now,	Comma indicates a level, continuing intonation; suggesting non-finality.
bu- but	A single dash indicates a cut-off (often with a glottal stop); including truncated intonation units.
DAMN	Capitals show heavy stress or indicate that speech is louder than surrounding discourse.
°dearest°	Utterances spoken more softly than the surrounding discourse are framed by degree signs.
says "oh"	Double quotes mark speech set off by a shift in the speaker's voice.
(2.0)	Numbers in parentheses indicate timed pauses.
	If the duration of the pauses is not crucial and not timed:
..	a truncated ellipsis is used to indicate pauses of one-half second or less.
...	An ellipsis is used to indicate a pause of more than a half-second.

ha:rd	The colon indicates the prolonging of the prior sound or syllable.
<no way>	Angle brackets pointing outward denote words or phrases that are spoken more slowly than the surrounding discourse.
>watch out<	Angle brackets pointing inward indicate words or phrases spoken more quickly than surrounding discourse.
[and so-] [WHY] her?	Square brackets on successive lines mark beginning and end of overlapping talk; multiple overlap is marked by aligning the brackets.
and= =then	Equal signs on successive lines show latching between turns of different speakers; they can also indicate that the turn of one speaker continues after e.g. backchannels of interlocutors.
H	Clearly audible breath sounds are indicated with a capital H.
.h	Inhalations are denoted with a period, followed by a small h. Longer inhalations are depicted by .h: (cf. prolonging colon above)
h	Exhalations are denoted with a small h (without a preceding period). A longer exhalation is denoted by h: (cf. prolonging colon above)
.t	Alveolar suction click
()	In the case that utterances cannot be transcribed with certainty empty parentheses are employed
(hard work)	If there is a likely interpretation, the questionable words appear within the parentheses.
/ /	slashes are used for phonemic transcriptions
((laughs))	Aspects of the utterance, such as whispers, coughing, and laughter, are indicated with double parentheses.
{waves}	Nonverbal behaviour, such as gestures, movements and looks, are indicated with curly brackets – for a more detailed overview of how to transcribe nonverbal behaviour see the separate taxonomy (http://umwelt-campus.de/case/conventions)
Numbering	Number each intonation unit consecutively (e.g. from 1 to n).

Conventions adapted from:

- Richard A. Dressler; and Roger J. Kreuz. 2000. "Transcribing oral discourse: A survey and model system." *Discourse Processes* 29. 25-36.
- Chafe, Wallace. 1982. "Integration and involvement in speaking, writing, and oral literature." *Spoken and written language: Advances in discourse processes*, vol. 9, ed. by Deborah Tannen, pp. 35-54. Norwood: Ablex.
- Chafe, Wallace. 1994. *Discourse, consciousness and time*. Chicago: University of Chicago Press.

Specific Transcription Conventions

For CASE there are some additional specific transcription conventions reflecting the particular nature and medium of discourse. They are as follows.

Start every transcript with the name of the recording, press enter twice and then start your transcription.

LANGUAGE STANDARD

For the purpose of consistency and to avoid complicated search algorithms, we will only use **British** standard in our transcripts. (Please change your spell-checking options to English UK.)

All lexicalized items are transcribed. You can use traditional (online) dictionaries and the Urban Dictionary. Conventionalized contractions such as *I'm*, *you're*, *it's* are transcribed as they occur.

PUNCTUATION

There always has to be some kind of intonation **punctuation** at the *end* of the last *spoken* word of *every* intonation unit to indicate rising (?) , falling (.) or continuing (,) intonation.

If aspects of the utterance occur **at the end** of an intonation unit, the punctuation precedes the feature.

okay. ((laughs))

If aspects of the utterance occur **while speaking**, the punctuation follows the feature, not the last word:

okay ((laughing)).

PAUSES

Commas are not to be used within intonation units, but only at the very end of them, to indicate a continuing intonation, as described in Dressler and Kreuz 2000.

Pauses between intonation units (n.x (tenths of a second), .. and ...) will begin the next intonation unit, not stand at the end of the preceding one.

CAPITALISATION

The **only** capitalised words are the first person pronoun **I**, **names**, and **nationalities**. Words or phrases that are spoken with a **higher volume** than the surrounding utterance are written in capital letters. The same applies to **especially stressed syllables**. (Do not capitalise regular stressing patterns.)

LAUGHTER

The following nine laughter types are distinguished in CASE:

((LAUGHS))	loud laughter (louder than surrounding discourse)
((laughs))	laugh pulse sequence
((laughing))	laughing while speaking
((chuckles))	chuckling, softer than laughter
((chuckling))	chuckling while speaking
((hehe))	short chuckling, two pulses
((thh))	aspirated minimum laughter starting with alveolar plosive sound
((ehh))	monosyllabic laugh pulse
((heh))	initially aspirated monosyllabic laugh pulse

ECHO

We will just transcribe echoes which are important for the discourse, i.e. which explain pauses/confusion because of speech delay, self-interruption etc.

€€€	normal echo
[€€€]	overlapping echo

Echoes are either transcribed as completely overlapping or not at all overlapping, i.e. there are no other transcription possibilities than the two above.

Each Echo is transcribed in a **separate line** as it occurs in the transcript, without indicating speaker provenance.

TECHNICAL INTERFERENCE (PACKET LOSS)

Technical interference will not be transcribed unless extensive and influencing the conversation. In contrast, stuttering and hesitation will be transcribed according to the guidelines.

If the recording is interrupted, please transcribe the interruption as follows:

#interruption of recording (3.7)#

CODE-SWITCHING

Code-switching will be transcribed.

If the speakers switch between two languages, indicate the language and the time (tenths of a second) this language is spoken: eg. **((Korean (3.4)))**

If you understand the language, include the transcription of it in your transcript.

PHONEMIC TRANSCRIPTIONS

We will not transcribe phonemically, except for mistakes in pronunciation, which might lead to confusion. First, write the word that was intended, then, indicate this particular pronunciation by placing **slashes** around the differing sounds and put it in **double brackets**. Example:

(1) SF13: so because I thought **((/t/ought))** if-

SF13 wanted to say “thought” but didn’t pronounce the <th> with its regular phoneme /θ/.

NUMBERS

Numbers (also years) will be spelled out: **two thousand (and) seven** instead of 2007.

OVERLAPS

[Overlaps] can happen in between or within words.

If two speakers overlap, but one of the speakers has a hesitation or pause within his turn, while the other one continues speaking, **group these overlaps to one single overlap**:

(2) SB14: YES,
ah sorry: for not informing you before ((laughing)),
but I've started the re[cording right-
... °right in the beginning]°.
SF13: [ye-eah it's- it's nice to do that],
because I don't wanna say something which is not a-ppropriate,

The intonation sign is after the square bracket. Don't align overlap.

If one of the speakers continues his intonation unit after the overlap, make sure that the discontinued speaker has two turns.

SB01: ... okay,
... my recorder is running automatically,
so I don't need to press anything,
SF01: [uh huh],

SB01: [it just] starts with the Skype thingy,

SF01: **okay mine does not do,**

BACKCHANNELING

There are no specific symbols to mark backchanneling.

CLIPPING

Words should not be shortened in any way. They have to be transcribed exactly as they are spoken.

Okay cannot be spelled “~~ok~~”.

LIAISON

If a speaker jointly pronounces several words and/or non-lexical items in an unusual manner, indicate by joining the items with underscores:

She said_uh

UNDERSTANDING PROBLEMS

If you cannot understand what has been said use normal, empty brackets without space in between: ()

If you think you understood something but are not quite sure, write these parts within the brackets: (example xy)

Do not specifically mention that you could not understand it, as the use of empty brackets already indicates this.

CONTOURS

Contours will not be included at this point. The symbols / and \ should not be used in a transcript, except for phonemic transcriptions / /.

BRACKETS

Normal, single brackets are used for understanding problems (see above).

Normal, double brackets are used to indicate aspects of the utterance.

Curly brackets/braces are used to indicate nonverbal behaviour.

Square brackets are used to indicate overlapping.

Do not use these brackets for any other purpose than the ones mentioned above.

INDENTS

Indents have the function of structuring the F4 transcript into a word table.

The general rule is:

NAME: ---press tab once--- TEXT

or

NAME: ---press tab once--- TEXT ---press tab once--- TIME STAMP

UNDERSTANDING CHECKS and HESITATION

mhm	closed mouth, agreement or understanding
m/m:/mh/m:h	closed mouth, single sound
uh huh	open mouth
uh/u:h	open mouth, single sound, often with hesitations
e:r	open mouth, clear pronunciation or /r/, often with hesitations
uhm/u:hm	open mouth that is closed at the end, also with hesitations

NAMES

Use the University/Student coding (e.g. SB01) for the speakers, also if they refer to each other. First and last names of other participants, co-students, university staff, friends or family of participants, and other names related to the personal background of the speakers (street address, city district, but not city name, study subject, faculty etc.) will be anonymized later – please mark them in the transcript with “!” after the word like this: Mozartstrasse!

All other names (cities, regions, countries, names of celebrities, TV series, movies etc.) will be transcribed in full.

CONCISENESS

Be concise and consistent. This means:

Do not use different spellings for the same word.

Check unfamiliar words. (The game *Counter-Strike*, for example, is neither spelled Counterstrike nor Counter Strike, and especially not Contra Strike).

Transcribing a mispronunciation and then using it as a normal word is not possible.

Please use the following citation:

- CASE transcription conventions (2017). Birkenfeld: Trier University of Applied Sciences.
[<http://umwelt-campus.de/case-conventions>]

A taxonomy for non-verbal elements in CASE (as of June 2017)

Please use the following citation if you would like to use it for your own work:

CASE Taxonomy for non-verbal elements. (May 2017). Birkenfeld: Trier University of Applied Sciences. [<http://umwelt-campus.de/case-conventions>]

Approach:

The CASE annotation scheme tries to refrain, as far as possible, from interpreting during annotation in order to make the transcripts as objective as possible, leaving it to the researchers to draw their own conclusions. The video component of CASE can be used to supplement the transcript for a more exhaustive multimodal interpretation, for example in the context of qualitative analysis.

At this point, there is no separate layer for non-verbal elements (NVEs) in CASE, but selected instances of non-verbal behaviour are included as part of the basic CASE transcription layer (**using curly brackets, e.g. {shrugs}**). Our use of the term **NVE** in the following includes gestures, as well as other elements that may occur, such as salient **head movements and facial expressions, gaze, physical stance shifts, actions and third-party interaction happening in interlocutor's surroundings, and background noise**.

Through a **data-driven approach**, we have classified the salient non-verbal elements transcribed in CASE and developed the taxonomy of NVEs described here.

General principles:

- Only **salient NVEs** are transcribed, i.e. NVEs contributing to / supporting meaning-making, as well as NVEs that are referred to on a verbal level or that refer to sth that is discussed
- Transcription uses **curly brackets**; as with all elements, brackets can be added outside the curly brackets to indicate unclear NVEs or instances when the transcriber is not sure how to transcribe the NVE correctly ({waves hands in the air and jumps up and down})
- The transcription should be **verb-based** and **concise**, that means if there are verbs that already encompass the NVE they should be used instead of disassembling the NVE into its single components,¹ e.g. {makes peace sign}
- If the NVE consists of separate stages that could be considered part of the same NVE, it should not be disassembled into phases but be **transcribed concisely** (e.g. {imitates breathalyzer by blowing into top end of pen}, not {imitates breathalyzer by lifting pen to mouth and blowing into top end of pen and setting it down}); however, various consecutive NVEs can be transcribed consecutively if no concise transcription exists:

¹ It is understood that this always will include some level of abstraction/interpretation - we do not aim at a semiotically precise representation

{smiles} {nods} {makes thumbs-up gesture}; if several NVEs co-occur, they are transcribed in one bracket and connected with &: {smiles & nods}

- NVE transcription is **not aligned or marked for duration, for intensity, or speed**
- Transcription should **follow** the intonation unit containing the most salient use (after intonation sign), or, if limited to smaller units (e.g. words), follow those units
- The **verb** should be in the **third person present tense**, not the present participle, i.e. **{nods}**, not {nodding}
- The following modifiers can be added (if co-occurring, in this order):
 - a) **direct object**, if the main verb does not already contain the object sense (e.g. “nods” includes the object “head” and does not need to be repeated), e.g. “lifts arm”, “lifts hand.”² If a NVE is already conventionally named, it may be used as object of “make”, e.g. {makes throwaway gesture} {makes peace sign} {makes air quotes}, further explained below
 - b) **temporal adverb** (three times, repeatedly, ...)
 - c) **directional adverb(ial)** (e.g. up, down, behind ear), if necessary. Directions are not separately denoted if the specification does not have an influence on the meaning of a NVE - thus, left and right are usually not distinguished. Directions are always given from the speaker’s perspective (outward, inward, front, back etc.).
 - d) **to ...**, if a target/level needs to be added (to window, chair, to eye, etc.)
 - e) **with ...**, indicating for example the hand(s), body parts or objects used in the NVE (e.g. with left hand, with left index finger), if salient (e.g. a thumbs-up NVE would not need a hand specification, as it is salient with either, while a both-hands thumbs up NVE would be more emphatic and warrant the modification “with both hands”
 - f) **by ...-ing**, if a further modification is needed (e.g. in the case of imitating, e.g. imitates breathalyzer by blowing into top end of pen)
 - g) other additions not covered, e.g. relation between elements or if further specification is needed, separated with semicolon {lifts hands; palms outward} {lifts hand; palm up} > this level of modification should be used sparingly

This sequence can be expanded to include several elements (but see above for the conciseness principle)

- In general, the taxonomy makes use of names of conventionalized Western European NVEs as descriptions to reduce complexity in the annotation; additional explanations are added in the transcription conventions, clarifying NVEs that might be culturally specific
- **NB:** The annotation aims to be as descriptive as possible. In some cases, the terminology used sometimes may imply certain meanings or interpret NVEs to a certain degree to clarify the context. The actual interpretation of the individual functions of a NVE, however, still lies with the researcher, particularly as there are different cultures involved which increases the probability of diverging functions for similar NVEs.

² here further modification may be needed, see below

Examples for CASE Taxonomy of salient non-verbal elements based on bottom-up analysis of CASE data (with possible additions and explanations; open-ended, can be completed acc. to conventions above as needed)

Head gestures, including gaze

- Looks (e.g. up, down, to side, to upper corner ...)
- Nods (i.e. head moves up and down)
- Shakes head (i.e. head turns left and right)
- Tilts head (repeatedly)
- ...

Facial expressions

- Frowns
- Grimaces (i.e. facial expression that usually implies a negative connotation such as distaste, revulsion, etc.)
- Purses lips (to side)
- Raises eyebrow(s)
- Rolls eyes
- Smiles
- Squints
- Winks
- Yawns
- ...

Hands / Body

- Claps
- Clasps hands (in front of chest if not otherwise specified)
- Drinks from ...
- Drums fingers (i.e. makes rapid alternating downward movements with fingers of one hand - except thumb - similar to playing the piano on solid surface)
- Eats ...
- Folds arms
- Hits/thumps with ... on ...
- Holds ... to ...
- Holds up ... (e.g. two fingers, glass of wine, etc.)
- Imitates ... (e.g. drinking, breathalyzer, braces, shape of ..., size of ..., etc., where necessary specifying with by ...)
- Lifts hand (e.g. to..., or lifts hand; palm up (often with explanations, cf. McNeill 1992))
- Makes ...
 - air quotes (i.e. imitates quotation marks with index and middle fingers of both hands)

- beat gesture (i.e. hand movement, usually up and down, during speech, often highlighting information, e.g. to signal significance (McNeill 1992, Wang & Chu 2013))
- box gesture (cf. Cassell 1998; i.e. raises both hands and/or moves them up/down, palms vertical, usually indicating a concrete object)
- fist pump (i.e. makes fist with one hand, moves fist quickly downwards and back towards body)
- fist(s)
- okay sign (i.e. puts index finger and thumb together with other fingers extended above index finger)
- peace sign (i.e. makes a 'V' with index and middle finger, palm outward)
- swiping gesture (i.e. moves hand sideways quickly)
- throwing-away gesture (cf. Bressemer & Müller 2014; i.e. makes quick downward movement with one hand, palm downwards)
- brushing-away gesture (cf. Bressemer & Müller 2014; i.e. makes quick upward movement with one hand, palm downwards)
- thumbs-up gesture, thumbs-down gesture (i.e. makes fist with thumb extended, pointing up/down)
- Moves hand to ... (e.g. to mouth, to forehead, etc.)
- Moves hands ... (e.g. in circle, outwards, up, etc.)
- Opens hand(s) (i.e. outwards movement, starting with closed hand(s), flipping wrist back, opening palm, ending with palm upwards)
- Points to ... (with ...) (e.g. with index finger/hand, etc.)
- Puts ... on ...
- Rubs ... against ... (e.g. rubs thumb against index and middle fingers; rubs index finger against underside of nose)
- Scratches (head) (if salient, e.g. in combination with hesitation, thinking, etc.)
- Shifts camera to show ...
- Shows ... (i.e. moves object in front of camera/close(r) to screen to focus attention on it)
- Shrugs
- Stands up / sits down
- Touches ... (e.g. head, ear, shoulder, etc.)
- Types
- Walks to ...
- Waves
- ...

Physical stance

- Leans ... (forward, backward, towards ...)
- Sits up (i.e. straighter than before)
- Shifts position
- ...

Background

- Movement: Noun + verb in third pers. through screen/towards/in front of screen ... (e.g. roommate walks through screen, roommate moves in front of screen, etc.)
- Noises: Noun + verb in third pers. if origin is clear (e.g. cat meows, baby cries, etc.), description as present participle or noun + sound (e.g. clicking sound, static sound, etc.)

Overview CASE transcription conventions (as of February 2017)

- **British** English spelling
- **Intonation punctuation** always at the *end* of the last *spoken* word of every intonation unit
- “xxx” and “nnn” = utterance

.	Falling tone in the preceding element; suggesting finality
?	Rising tone in the preceding element
,	Continuing intonation; suggesting non-finality
-	Cut-off (often with a glottal stop); including truncated intonation units
CAPS	Heavy stress or loudness, can affect syllables or whole words. The only capitalised words are the first person pronoun I, names, and nationalities.
°xxx°	Utterances spoken more softly than the surrounding discourse
"xxx"	Speech set off by a shift in the speaker's voice
(n.x)	Timed pauses (n=seconds, x=tenths of a second)
..	Pauses of one-half second or less
...	Pause of more than a half-second
:	Prolonging of the prior sound or syllable.
<xxx>	Words or phrases that are spoken more slowly than the surrounding discourse
>xxx<	Words or phrases spoken more quickly than surrounding discourse.
A: [xxx], B: [nnn],	Beginning and end of overlapping talk (on successive lines), in between or within words
A: xxx= B: =nnn	Latching between turns of different speakers (on successive lines)
H	Clearly audible breath sounds
.h	Inhalation. Longer inhalation depicted by .h: (cf. prolonging colon above)
h	Exhalation. Longer exhalation denoted by h: (cf. prolonging colon above)
.t	Alveolar suction click
()	Utterances that cannot be transcribed with certainty
(xxx)	Likely interpretation of uncertain utterances
((xxx))	Aspects of the utterance, such as whispers, coughing, and laughter
((/xxx/))	Phonemic transcriptions. Only problematic cases, e.g. “thought ((/t/ought))”?
((Language (n.x)))	Code-switching. Indicate language and duration (n=seconds, x=tenths of a second), cf. timed pauses
€€€	Normal echo, transcribed in a separate line
[€€€]	Overlapping echo, transcribed in a separate line
{xxx}	Nonverbal behaviour, such as movements and looks

#interruption of recording (n.x)#	Interruption of recording and duration (n=seconds, x=tenths of a second), cf. timed pauses
xxx_nnn	Liaison, if unusual
number/year	Spelled out

Other features:

((LAUGHS))	Separate, loud laughter
((laughs))	Separate, normal laughter
((laughing))	Laughing while speaking (and other laughter types uttered while speaking), belongs directly to the word, followed by intonation punctuation
((thh))	Aspirated/minimum laughter
((chuckles))	Chuckle, softer than laughter
((hehe))	Short chuckling, two laugh pulses
((heh))	Short (monosyllabic) laughter (one laugh pulse), initially aspirated
((ehh))	Short (monosyllabic), exhaling laughter, initially not aspirated
((xxx-ing))	While speaking, cf. laughing
mhm	Closed mouth, agreement or understanding
m/m:/mh/m:h	Closed mouth, single sound
uh huh	Open mouth
uh/u:h	Open mouth, single sound, often with hesitations
e:r	Open mouth, clear pronunciation or /r/, often with hesitations
uhm/u:hm	Open mouth that is closed at the end, also with hesitations