

Patterns and Variation in English Language Discourse

**9th Brno Conference
on Linguistics Studies
in English**

edited by
Irena Hůlková
Renata Povolná
Radek Vogel

**MASARYK
UNIVERSITY
PRESS**

**Patterns and Variation
in English Language Discourse**

9th Brno Conference on Linguistics Studies in English

Irena Hůlková, Renata Povolná, Radek Vogel (eds)

Masaryk University Press

Brno 2022

Reviewed by: doc. Mgr. Martin Adam, Ph.D.
Mgr. Christopher Hopkinson, Ph.D.



Kniha je šířená pod licencí
CC BY-NC-ND 4.0 Creative Commons Attribution-NonCommercial-NoDerivatives 4.0

© 2022 Masaryk University
ISBN 978-80-280-0212-1 (online ; pdf)
ISBN 978-80-280-0211-4 (paperback)
<https://doi.org/10.5817/CZ.MUNI.P280-0212-2022>

CONTENT

Introduction	5
The multifaceted and whimsical nature of discourse	7
<i>Jarmila Tárnayiková</i>	
Repetition as an effective rhetorical device in remarks at media briefings on COVID-19	26
<i>Zhuzhuna Gumbaridze</i>	
Tendencies of threat techniques utilization in the COVID-19 related discourse in Georgian and English online news reports	42
<i>Zeinab Gvarishvili and Nana Mazmishvili</i>	
Adverse ‘native speaker’ effects on Anglophones in the multilingual workplace	59
<i>Veronika Lovrits</i>	
Hedging like a professional: A corpus-driven approach to interactional metadiscourse in English learner academic writing	72
<i>Markéta Malá</i>	
A comparison of climate change lexical creativity among American, European and international Twitter users	87
<i>Vanessa Marcella</i>	
Possibility modals in English tourism discourse: Variation across three web registers.....	105
<i>Aleksandra Radovanović</i>	
Grammatical collocations in English exam texts written by Czech secondary-school students	120
<i>Věra Sládková</i>	
Aspects of vocabulary knowledge in current lower-secondary EFL coursebooks and coursebooks published 20 years ago	142
<i>Michaela Trnová</i>	

PATTERNS AND VARIATION IN ENGLISH LANGUAGE DISCOURSE

INTRODUCTION

The current volume presents a selection of articles written by participants of the Ninth Brno Conference on Linguistics Studies in English entitled *Patterns and Variation in English Language Discourse*. The conference was organised by the Department of English Language and Literature at the Faculty of Education of Masaryk University in Brno and was held on 16–17 September 2021. The event hosted more than 50 participants from different countries and attempted to extend our understanding of language as a social, contextually bound phenomenon while exploring how patterning and variation in texts and discourses can be shaped by the communicative needs, the topic at hand, the medium employed and the intended audience.

The papers in this book include one of the three plenary speeches (delivered by Prof. Jarmila Tárnyiková from Palacký University Olomouc) and another eight contributions, all of which map new trends in the study of discourse, characterised by the alliance of corpus and discourse analysis, multimedia studies, intercultural rhetoric and cross-linguistic studies.

Jarmila Tárnyiková explains how the dynamism of human evolution and the consequent changes in communicative strategies have enhanced a whole spectrum of new challenges evoked by changes in the scope of items considered relevant for present-day research and above all by the existence of language corpora, which offer an immensity of data across genres, language varieties and language interfaces. Her aim is to map the facets which contribute to patterning and variation in discourse shaping, and by using authentic language data amplify the role of peripheral language devices in interaction.

The paper by Zhuzhuna Gumbaridze looks into the use of repetition in speeches delivered at the time of the COVID-19 pandemic by two top representatives of the World Health Organization and a Georgian national disease control authority. She claims that repetition is employed as one of the principal rhetorical devices with the aim of enhancing credibility and exercising a direct impact on the audience.

Public discourse in the time of the pandemic is also the focus of research by Zeinab Gvarishvili and Nana Mazmishvili, who have researched online news in English and Georgian and established different densities of citing authorities, citing patients, using metaphors, informing about objective threats and other threatening and persuasive strategies in the two corpora.

Veronika Lovrits' contribution explores the social effects of the differentiation between 'native' and 'non-native' varieties of English. Her sociolinguistic analysis of stances – while discussing participants' reflections on language practices at work and mapping their discursive positioning – points out the social constructivist character of the native/non-native dichotomy and refers to its adverse effects in the multilingual workplace and beyond.

Markéta Malá looks into the phraseology of English academic texts written by Czech university students and compares these with English L1 novice and expert academic writing while focusing on hedging patterns. Her findings show that in terms of the use of epistemic stance in their English academic articles, Czech university students have to deal with a lack of academic experience and – regardless of their proficiency – also a lack of linguistic EFL practice.

Vanessa Marcella presents a case study analysing lexical innovations in climate change discourse in the micro-blogging service Twitter with a focus on the use of the English language through hashtags, and concludes that by means of the creativity of hashtags, Twitter users can contribute to a valuable interpretive framework.

Aleksandra Radovanović focuses on modal verbs of possibility used in tourism discourse. Her analysis involves three genres, namely promotional texts, electronic newsletters and managerial responses to guests' reviews. The use of modals displays clear situationally-determined differences, with functions ranging from conveyance of information to interactivity and permission.

Věra Sládková studies the frequency and accuracy of five types of G8E–G8I grammatical collocations (Benson et al. 1986) in an English language learner corpus (CZEMATELC) and reveals the prevalence of A1–A2 CEFR level colligations relying on a limited number of verb lemmas, a wide incorrect pattern variation and preference for patterns which are also the most frequent patterns of their Czech equivalents.

Finally, Michaela Trnová presents a content analysis of vocabulary activities in selected lower-secondary coursebooks used in the Czech Republic with the aim of verifying the categorial system based on Nation's (2001) nine aspects of word knowledge as an appropriate research tool for a comparison of approaches to vocabulary in the current teaching materials and in those published 20 years ago.

We would like to express our great appreciation and thanks to all the authors for their contributions and the reviewers for their constructive and valuable comments. We believe that the present volume will serve as an incentive to further exploration for anyone interested in linguistics studies in English.

The editors

THE MULTIFACETED AND WHIMSICAL NATURE OF DISCOURSE

Jarmila Tárnyiková

*...the greatest and most abiding difficulty about
human communication is that it is human.*

Quirk (1987: 9)

<https://doi.org/10.5817/CZ.MUNI.P280-0212-2022-1>

Abstract

My contribution, rooted in functional and systemic grammar, is based on the assumption that though discourse as a social behaviour and verbal interaction has been studied by great minds for decades, the dynamism of human evolution and the consequent changes in communicative strategies can hardly leave discourse analysts immune to a whole spectrum of new challenges. These are evoked by changes in the scope of items considered to be relevant for present-day research, by blurring the traditional borderlines between categories (written manifestation of spoken discourse in chatting), but before all by the existence of language corpora offering the immensity of data across genres, language varieties and language interfaces. A brief introduction (Part 1) will be followed by three main parts, focusing on reasons for multifacetedness in discourse (Part 2), whimsical nature of discourse (Part 3), and hands-on experience with overt language manifestations of vagueness, as exemplified by English placeholders (*Mrs Thingy*, *John Whatsisname*, *whatchamacallit*, *so-and-so*) emergent from the BNC and COCA corpora (Part 4). The aim is twofold: to map the facets which contribute to patterning and variation in discourse shaping, and by using authentic language data amplify the role of peripheral language devices in interaction.

Keywords

discourse, text, facets, metafunctions, principles, vagueness, placeholders

1 Introduction

1.1 State of the art

Over the years, the study of discourse has undergone dramatic changes: new methods, new theories, case studies, interdisciplinary connections but also new journals, new book series, academic programs and a plethora of conference activities all over the world. All this reflects the dynamic and evolving nature of the field (cf. Biber et al. 2007).

With the central theme of the conference in mind, I opted for a topic that would echo both the processes of patterning in discourse and the existence of variations, opening a space

for multiple choices, multiple processes underlying language use, and multiple processes of language change, imposed on us by dynamic processes of evolution in the society we live in, by the globalized world we share, and by intensive language contacts.

1.2 Procedure

My contribution will proceed from a roadmap of perceiving the status of discourse – to a narrowed pathway of tracing selected discourse properties, with my role fittingly characterized by Fillmore (1992) as a computer-aided armchair linguist, who admires the immensity of corpus data but feels the responsibility of being qualified to benefit from the data by considerate armchair theorizing.

1.3 Keywords

All the keywords mentioned above have one in common: they relate to the level of language analysis that goes beyond the sentence and relates to the ways *utterances* (rather than sentences) work in sequences or stretches of language – spoken or written, short or long – not as self-contained units of meaning built *in a brick-by-brick fashion* but as essential parts of the meaning of their adjoining utterances, whether connected by grammar or simply by sequencing. They are shaped in a Rubik’s cube-like fashion as appropriate contributions to a given communicative situation (compare Fillmore’s 1976 *frame theory*). It is notorious that in our everyday interactions, the sentence is only a small cog in a much larger machine.

1.4 Aim

Rather than preferring one paradigm over all others, my aim is to defend the plurality of voices, in which each voice is perceived as a part of a chorus of other voices. In my view, there are many valuable ways of approaching discourse, any of which may prove useful or stimulating. But, whichever voice is presented, advocated, or criticized, the conception should be adequately presented, i.e. interpreted in a way sensitive to the respective theoretical framework in which it is anchored, and in its complexity; only then it can be critically assessed.

To exemplify the case, let me remind you of Halliday’s (1979) *constituents* (ideational, interpersonal and textual), i.e. the term used in his earlier studies, and his term *metafunctions*, prototypically used in his later studies. Both refer to comparable cognitive entities, looked upon, however, from different perspectives (bottom-up in case of constituents,

and top-down in case of metafunctions, see Halliday & Webster 2009: 6, 253). Moreover, in his architecture of language (Halliday 2003), the terms are incorporated into networks of language systems based on different scales of delicacy. So, when discussing Halliday's conceptions and taxonomies, we should not ignore the appropriate timing of their origin.

Similarly careful thoughts should accompany our theorizing on *speech acts*. The initial and rather synthetic view of speech acts, as reflected in Austin's (1962) and Searle's (1969, 1975) taxonomies, has been gradually giving way to analytical, *speech-act-set* approaches offering a more delicate scale of scenarios, inspired by authentic corpus data. In Leech (2014) these are referred to as *speech act territories*.

The samples below, illustrate two- and three-member sets of speech acts emergent from Svartvik's (1990) *London-Lund Corpus* (LLC).

IFID + apologetic account

- (1) *I'm terribly sorry* [1] / *but I shan't be with you until five past ten.* [2] (LLC, 130),
in which IFID stands for *illocutionary force identifier* (prototypically *I'm sorry*).

IFID + apologetic account + disarmer

- (2) *I'm sorry* [1] / *about the mess,* [2] / *how stupid of me.* [3] (LLC, 281)
(for a survey of more elaborate scenarios compare, cf. Tárnyiková & Válková 2000)

There is a fair measure of agreement among linguists about the importance of speech-act-sets in cross-language comparison. Speech-act sets help identify similarities, differences and preferences in the etiquette of the compared culture-bound language communities. Thus, for example in Czech, the bus driver, when out of service, informs the potential passengers by a pragmatically predictable statement *Mimo provoz*, while the English bus driver displays visibly *Sorry, I'm out of service*, with the institutional apology worded as if it were a personal apology of the driver.

1.5 Theoretical framework

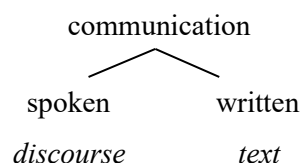
As a functional linguist rooted in the Prague school of functional and systemic linguistics, I approach language as a system of systems and a pattern of patterns. Mindful of the scalarity of central and peripheral means of language expressions, I avoid the strategy of sweeping problematic or peripheral cases under the carpet, since, as the corpus data reveal, many of the peripheral devices can be endowed with special communicative roles (cf. informal means of vague language discussed in Section 4).

1.6 A toolkit of prerequisites

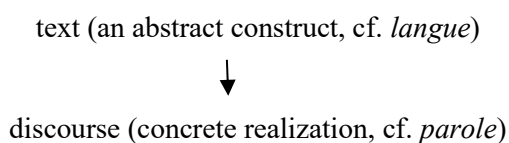
In order to share a manageable framework of reference, let us begin with recapitulating some key terms and processes associated with discourse analysis. Before doing so, however, we have to consider the technical question of *what kind of interaction we are talking about*, since one of the challenges discourse analysts are faced with nowadays is how to cope with the *masked interaction* imposed on us by the novel coronavirus pandemic, in which the face-to-face interaction has given way to a mask-to-mask interaction, neutralizing in many respects the discourse-dove and discourse hawk strategies. So, is it still a facework or a face-less facework? Well, let's hope it will be a bad joke in the near future.

1.6.1 Discourse and/or text

The two terms, *discourse* and *text*, are used in a very inconsistent way in literature, thus reflecting various conceptions, degrees of abstraction, types of hierarchy, etc. Many scholars associate *text* with written communication and *discourse* with spoken interaction, as visualized below.



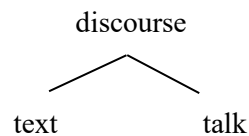
For others (Enkvist 1994), *text* is an abstract construct (similar in function to *langue*), while discourse is its *parole* manifestation.



With the occurrence of *Text Linguistics* in the 1970s, and the search for *parameters of textuality* or *textness*, there was a strong tendency with text analysts to make a distinction between *text* as a neatly woven texture (with the focus on such text parameters as *cohesion*, *coherence*, *intentionality*, *acceptability*, *informativity*, *situationality* and *intertextuality*) and *discourse*, primarily reserved for spoken data and the mechanisms of conversation manoeuvres, such as turn-taking, turn-passing practices, the study of back-channel

communication; strategies of ongoing checks or changing topic, the role of *adjacency pairs* in conversation, *conversational routine* and many more (Schegloff 1982).

A diplomatic solution was proposed by Deborah Tannen (1982). In her view there are two separate *modes* of discourse and, consequently, the umbrella term is *discourse*, subcategorized into *text* (written prose) and *talk* (spoken conversation), i.e.



What Tannen (ibid.) called two *separate* modes will probably nowadays be referred to as two *interlinked* modes of discourse, representing a scale rather than a dichotomy (cf. chatting, e-mails, or, for example, spoken passages in written texts simulating spontaneous face-to-face interaction). What seems to prevail nowadays is the term *discourse* used as an umbrella term for both spoken and written manifestations of language use. As Tannen (1982: x) put it, “It is fortunate ... that there exists in English a word that refers to language in context across all forms and modes. That word is discourse”.

1.6.2 Discourse analysis (DA)

Though very often approached as a separate linguistic discipline, discourse analysis is a *research method* for studying written or spoken language in relation to its social context. It is a method of analysis cutting *vertically* across all the levels of language representation (phonetics, morphology, syntax ...) and encapsulating both communicatively constitutive and communicatively regulative units of interaction (to be discussed later).

The scope of DA can be either very local (conversation of two people, or a single diary), or very global and abstract.

2 Multifaceted nature of discourse

The term *facet* (known, for example, from crystallography, cf. a multifaceted cut of a diamond), is used metaphorically here to refer to a particular aspect, perspective, dimension or feature of discourse (Compare also the metaphoric use in current use, as in *many facets of her personality*, *many facets of the problem*.).

Our starting point will be Halliday’s metafunctions already mentioned in Introduction. In the sections below, I will gradually extend the survey by other contributions to

multifacetedness, comprising both obligatory and optional facets. The aim is to visualize how many facets we have to control in trying to achieve our intended communicative goals and how the facets operate in a complex interplay.

2.1 The existence of metafunctions

Echoing Halliday and Webster (2009), we can state that discourse is a manifestation of three grammatically relevant *metafunctions* of the semantic system of language: *ideational*, *interpersonal*, and *textual*. Since Halliday's conception is generally shared by linguists, it would be redundant to go into details here. Put simply, we have mechanisms in natural languages how to convey meaningful pieces of information, how to add attitudes to our messages, and how to organize our messages as coherent and cohesive contribution to the surrounding context (both explicit or implicitly retrievable from a given communicative situation). Overt language means manifesting the three facets are isolable but operate in a complex interplay. The question is whether there is a hierarchy in their roles. In Enkvist's (1994) view, the text metafunction is superordinate, but his standpoint is not generally shared. What is, however, shared is the possibility to subcategorize within each of the metafunctions (compare Halliday's 1979 initial step in subcategorizing ideational component into *experiential* and *logical*).

Emergent from recent studies on discourse properties is an increased interest in subcategorizing interpersonal and textual metafunctions and projecting the means of their manifestation as *scales* rather than isolated *entities* (compare the scales of modal, evaluative, or emotional attitudes, ranging e.g. in modality from certainty via probability and possibility – to uncertainty).

2.2 The existence of social dimension and principles of language use

In considering these *social* facets, I base my arguments on Halliday's (1978) conception of language as social semiotic, with *man as social man* born in a particular language community and sharing with that community rituals, values and culture-bound preferences in language use. The social dimension is projected into discourse by strategies in cooperation and politeness, based on Grice's (1975) Cooperative Principle, with the maxims of *quality*, *quantity*, *relevance* and *manner*, and Leech's (1983) Politeness Principle, with the maxims of *tact*, *generosity*, *approbation*, *modesty*, *agreement* and *sympathy*. Both the principles are associated with the sociolinguistic notion of *facework*, as proposed by Goffman

(1956). Unlike the facets based on the existence of metafunctions, here the facets are based on principles, subcategorized into maxims.

Associated with the social dimension is also the *impact of language contact*. In the globalized world, the intensity and frequency of language contacts is beyond any doubt. The dynamism of the process of borrowing can be traced in two stages, referred to as contact-induced language *choice* (when the loanwords co-exist with the words of a domestic word stock as possible alternatives), and contact-induced language *change* (when the loanword replaces the domestic expression or compensates for the absence of a comparable equivalent in the target language). The speed of the process depends on the communicative priorities of the recipient language community and the need to name new entities, processes and events emergent from the globalizing world (for details, cf. Tárnayiková 2009).

The dynamism of choices and changes represents a social-bound variable. The following facets are associated with various kinds of perspectivization, layering or levelling in discourse analysis, and the interplay of verbal and non-verbal interaction.

2.3 The existence of pragmatic dimension

The facets associated with pragmatic dimension in language use are numerous, since pragmatics, contrary to those who approach it as a language discipline is perceived here with Verschueren (1999) as a *perspective*, cutting vertically across all the traditional levels of language representation (phonetics, morphology, syntax, ...) and finding means of manifestation in all of them.

2.4 The impact of the principle of language economy

Theoretically, its existence can be explained by Zipf's (1949) psychobiological *principle of least effort*, which, projected into language use, relates to man's inborn tendency to invest the least expenditure of effort to achieve the greatest amount of information within a given discourse. As a result, we are faced with partial or total reductions in discourse (cf. the reduction of a noun phrase, such as a *handsome man of forty* into pronominal *he*, or the elliptical structure *I would if I could* with the missing parts retrievable from the context).

Emergent from informal on-line communication are structurally reduced substandard lexemes and sentences, referred to as alphanumeonyms, such as *4U* ← *for you*; *GR8* ← *great*; or *3B* ← *bla, bla, bla*; *TBL* ← *Text back later*; *W84M* ← *Wait for me!*; *J4F* ← *just for fun*.

2.5 Potentiality of the phenomena of language

The notion of potentiality is known to Czech linguists since Vilém Mathesius' paper from 1911. Here, I would like to mention a type of language potentiality based on the existence of gaps in language taxonomies and the relative possibility to fill them. In doing so, we can achieve an attention-getting (and often humorous) effect in discourse.

Compare the gap in deriving feminine counterparts from masculine nouns, as in *lev* (lion) – *lvice* (lioness) in Czech, but the potential gap in *had* (snake) and **hadice*, in which the latter, if used, means *hose* we use in e.g. watering our garden.

With some adjectives, the polarity pairing is also missing in Czech, so if you fill the gap, you can achieve jocular results, as in *netečný* – **tečný*, cf. *Zatím co já jsem k pití netečný, ty jsi často až moc tečný* (literary: While I am indifferent to drinking, you can often drink quite a lot.). To exemplify the potentiality in English, we could mention the word-formative processes and the absence of infixation. Emergent from language corpora, however, is a tendency to fill the potential gap by using infixation as an intensifying device, as in *absolutely* > *absobloominlutely*, or *California* > *Calibloodyfornia*.

Appended to our tentative list of potential sources of multifacetedness can also be the scales of *explicitness* and *indirectness*, *intentional deviation* (*Heinz Meanz Beanz*), or *ambiguity*, opening space for more than one interpretation, as in *The police were ordered to stop drinking after midnight*.

2.6 The existence of two modes of discourse: Spoken and written

Spoken and written forms of discourse activate the same language code but have different hierarchies and preferences in variation and patterning, due to different targets to achieve. With the plethora of studies on written and spoken language, it would be redundant to discuss their mutual relationship, namely in the age of new technologies, in which the borderlines between the two tend to be blurred in some types of discourse, giving way to a variety of mixed forms (cf. written versions of spoken messages in on-line chatting).

Crystal (2001: 238) amply described the situation by stating that “in digitally-mediated communication, we are faced with neither spoken writing, nor written speech but with a totally new medium”.

2.7 Two types of language units: Constitutive and regulative

In analyzing discourse across types and varieties, we are faced with two types of language units. The distinction is not visible at first sight, namely if we take into consideration the principle of linearity in shaping our messages (especially in written discourse). Communicatively *constitutive units* create the core of our messages, and are *rule-governed*. Complementing them in various ways are communicatively *regulative units*, supplying the interlocutors with means used to signal the manifestation of conversation principles, as enumerated in Section 2.3. Unlike the rule-governed nature of constitutive units, the regulative units (discourse markers/signposts, hedges, mitigating devices, text-segmenting signals, face-saving formulae, and many more), are *principle-controlled* in nature.

While with constitutive units you can learn the rules, with regulative units you have to be socialized into the ‘dress-code’ of their use, taking into view social- and culture-bound preferences of the respective language community and the quantitative and qualitative parameters of their functional load.

As a result, we are faced with *two types of linearity* – one based on sequences of *homogeneous items* (e.g. the rule-bound sentence patterns, or morphological taxonomies) – and one, based on heterogeneity (a mixture of constitutive and regulative units, as in:

(3) *Well, as a matter of fact, Peter, to tell you the truth, is a lazy boy.*

2.8 The existence of segmental and suprasegmental levels

Though there seems to be a consensus among linguists that suprasegmental units of language are perceived as operating on segmental units, it is notorious that intonation contour, similarly to pitch, loudness, tempo, stress or voice quality, can become important decision-making cues in decoding messages.

In literary text, for example, we are often faced with suprasegmental *prompts* of those paralinguistic cues, bracketed in the main body of the text to give guidance to the intended interpretation (*she whispered, her voice was trembling, he heaved a sigh saying ...*). In small talk, the responses like *It's not what you are saying but how you are saying it.* – give us evidence of how changes in pragmatically shared intonation contours can influence the process of decoding.

2.9 The interplay of verbal and non-verbal forms of interaction

If we leave aside sign language, we can state that non-verbal forms of interaction are usually approached as supportive to verbal interaction, supplying language users with various interpretative cues. The diplomacy of human interaction, however, opens space for non-supportive use of non-verbal means (compare, e.g. the disdain, mockery or irony emergent from saying something while non-verbally signalling something opposite, as in *So, what was the dinner like? – Oh, it was delicious!* said while winking at the third party). I refer to the situation as shadowing of the facets, since we are saying something and meaning something else.

2.10 Summing up

If we project the identified facets into an imaginary matrix, we might perhaps state that the multifacetedness of discourse is caused by partly obligatory and partly optional configurations of heterogeneous facets (cf. metafunctions, principles and maxims, parameters, perspectives, scales or clines...) operating in a complex interplay, in proportions sensitive to the type of discourse and purpose to accomplish. Our camera-eye of shaping discourse can be partly *predictable* and partly *innovative*, and our perspectivization in discourse can be participant-oriented (as in diaries), event-oriented (reporting on sports events), time-oriented (chronicle), and place-oriented, as in guidebooks.

3 The whimsical nature of discourse

The whimsical manifestations of language in discourse are mostly associated with the intentional (and partly non-intentional) flouting of The Cooperative Principle and its maxims, as proposed by Grice (1975) and elaborated by Leech (1983) (for details, cf. 2.2). To exemplify the case of flouting: why do we use such phrases as *to tell you the truth*, or *to speak to the point*, when we are supposed to do so automatically? Why do we violate the maxim of quantity by using the whole chain of hedges by which to sandwich, as it were, the main body of the utterance, as in:

(4) *Well, as a matter of fact, the way I look at it, she is a sort of nice person or whatnot.*

In Robin Lakoff's (1982: 35) view "...flouting of the maxims is due to our desire to adhere to more socially (as opposed to intellectually) relevant rules, rules of Rapport: when

we have a choice between being offensive and being unclear, we invariably choose the latter, and a majority of cases of ordinary conversation can be seen to stem from this assumption”.

3.1 Strategies of being ‘unclear’

For being *unclear* we have a whole battery of strategies of intentional smoke-screening, wooliness, indeterminacy, and vagueness, contributing to a whole scale of *diplomatic manoeuvring* (typical e.g. of political speeches of these days), and a *whole battery of language devices* subsumed under the notion of *vague language* (Channell 1994).

3.2 Approaches to vagueness

Vagueness is not primarily approached as a *lack of sound argument*, or a result of a semio-chaotic behaviour, or the absence of coherent meaning (cf. the *sloppy* or *porous* communication, with which vagueness is often associated), but as a remarkably intricate part of social interaction (cf. the intersection of vagueness and politeness in facework strategies, i.e. not to lose one’s own face and not to threaten the face of the other).

There are countless communicative situations in which we are intentionally vague – for various reasons and with various communicative intentions in mind. The strategy of using vague language is so deeply rooted in our subconscious processes that we tend to take the language manifestations of vagueness for granted, as the flesh and blood of our everyday encounters. *A pile of papers can be twenty or so, the story has been told umpteenth times before, the playground is used by hords of unsupervised teenagers, so that in fact with oodles of patience we have bags of possibilities and stuff like that; in the shop, we tend to ask for a few apples and are happy when offered five; you can meet Mr Thingy or Ms Whatsername or whatnot.*

3.3 Benefits of vague language

Vague language is used to make the communication more informal and less tense. Vague language can contribute to vividness of our encounters by putting an additional flavour of expressivity and emotionality into our exchanges, as e.g. is the case of vague non-numerical quantifiers in *exaggerating* (*There were mountains of litter and heaps of directives in my office last week.*), *boasting* (*With his bags of talent my son must win.*), *complimenting* (*With your oceans of energy you can do miracles.*), or expressing *disgust or sarcasm* (*There is*

not a shred of truth in what you're saying. There is not an iota of common sense in his arguments.)

Some of the vague expressions can serve as therapeutic vague fillers used to fill a gap for temporary forgotten words or words that are at the tip of your tongue but not within linguistic reach, such as *Where is the thingy?*. These are referred to as placeholders and will be discussed in the following section (for types of vague reference, cf. Tárnayková 2009, 2010, 2019).

4 Hands-on experience with corpus data: English placeholders

What follows is a thumb-nail sketch of a research into one of the manifestations of vagueness, as represented by placeholders emergent from Mark Davies' BYU suite corpora (cf. References). The focus is on their formal, functional and distributional properties in discourse, and the aim is to prove the significance of peripheral language items in informal spoken interaction.

First, the working definition of placeholders: "Placeholder names are words that can refer to objects or people whose names are temporarily forgotten, irrelevant, or unknown in the context in which they are being discussed". (https://en.wikipedia.org/wiki/Placeholder_name).

In the next sections I hope to prove that the definition is simplistic, ignoring the double role of the placeholders, emergent from authentic corpus data.

4.1 Illustrative samples

Knowledge deficit

- (5) *Wait, Elliot, do you still have that thingy on your thingy?
– You mean GPS on my tablet?
I speak stupid. (BYU-EDU TV Corpus 84-2016 Young and Hungry)*
- (6) *Calvin, bring me the, uh...the-the...what...whatchamacallit.*
(BYU, TV Corpus, Nash Bridges, 46, 1997)
- (7) *Oh, he'll hang me from the highest, uh...Um, whatchamacallit...Uh, seven letters. Starts with "Y". – Yardarm. (M*A*S*H 67, 1977)*

Arbitrary use in legal discourse

- (8) *What is she, intoxicated or something?
Jane Doe, no I.D. She's confused and disoriented to the time and place.
(BYU, TV Corpus, Grey's Anatomy, 7, 2017)*

Chaining of placeholders

- (9) *It could be a thingamajig or a who's-he-what's-he or a Whatchamacallit.*
 (BYU-EDU, TV Corpus, The Office, 35, 2007)

Intentional dropping of the name

- (10) *Nah. soon's no good for me. I'm having lunch at the whatchamacallit with what's-his-name. Hi.Jack.Bye.* (BYU-EDU, TV Corpus, Will & Grace, 44, 1999)

Below is a table surveying the data. Most of the placeholders are borrowed from Channel (1994), some are based on my corpus findings. Generally speaking, we are faced with a partial grouping into TH-words, WH-words and G-words, and a set of individual entries mostly standing for placeholders of location (*Backwater, Podunk*, etc.).

Based on Channell's list	My addition
<i>thingy</i>	<i>whatchamacallit</i>
<i>thingummy</i>	<i>Joe Doe / Jane Doe</i>
<i>thingummyjig</i>	<i>so and so</i>
<i>thingummybob</i>	<i>you-know-who</i>
<i>whatsisname (called)</i>	<i>Backwater</i>
<i>whatsername (called)</i>	<i>The back of beyond</i>
<i>whatsitsname</i>	<i>Podunk</i>
<i>whosit</i>	<i>Timbuktu...</i>
<i>whatsit</i>	<i>gadget, gimmick, gizmo, widget</i>
<i>whatnot</i>	

Table 1: A survey of data samples

4.2 Status of placeholders and their locus

The status of PHs, compared to the status of the proposition in which they occur, is peripheral, and restricted to informal situations of spoken language use. There are, however, socio-cultural preferences with some kind of discourse, in which placeholders are obligatory, as evident from legal discourse:

- (11) *Investigators still have not been able to identify the missing suspect known as 'Joe Doe 2'.*
 (BYU-COCA 1995 SPOK ABC_DayOne)

Prototypically, placeholders occur in informal spoken language. There appeared to be socio-cultural preferences in their frequency of occurrence in particular types of discourse (cf. the already mentioned arbitrary use in legal discourse).

4.3 Problematic spelling

The following three types of variations occurred in the data:

- a) Fused (compact) forms, as in *Whatchacallit*, *whatsisname*, *whatsit*, *whatsits*, *whosis*, *whosit*;
- b) Hyphenated chains, e.g. *What-d'ya-call-it*, *Thing'em-bob*, *What's-his-name*, *You-know-who*, *So-and-so...*; and
- c) Non-hyphenated chains, as in *What d'ye call him*.

4.4 Word-formative processes

Emergent from the data were two word-formative processes:

- a) Squeezing, as in
What you may call it > *whatchamacallit*
What is it > *whatsit*
- b) Extension
thingy > *thingummy* > *thingumbob* > *thingumabob*

(Compare analogy with the extension of pronominal placeholders in Czech, as in *ten – tendle – tendleten*, *tendlencten – tentononc*.)

4.5 Data processing

The focus in my research was on *nominal* placeholders, as opposed to verbal placeholders (*We've whatchamacalit, haven't we?*) or adverbial placeholders (*He comes from the back of beyond*).

In analyzing corpus data, two axis of research have been activated: the *vertical*, *paradigmatic* axis, cutting vertically across the obtained data, which enabled me to trace possible alternatives within PHs (*thingy/whatsit/whachamacalit...*) and their frequency of occurrence, and a horizontal axis of mutual co-occurrence of the PHs and/or their co-existence with other regulative units in discourse, cf.

(12) *Er it's whatsit it's some er Norwegian thing* (BNC KR2)

(13) *I don't I don't whatsername for all.* (BNC FYG 986)

This enabled me to support my hypothesis about more than therapeutic functions of PHs. The following frequency surveys offered a good guidance to my goals.

Placeholders	BNC	COCA
<i>thingy</i>	131	81
<i>thingybob</i>	10	0
<i>thingummy</i>	8	8
<i>whatsit</i>	71	25
<i>whatsisname</i>	13	9
<i>whatchamacallit</i>	4	0
<i>whatchamacallem</i>	1	0
<i>whosit</i>	1	3
<i>so and so</i>	113	317
<i>Joe Doe</i>	1	5

Table 2: Frequency survey based on BNC and COCA corpora

Placeholders	BNC	COCA	TV corpus	Movie corpus
<i>thingy</i>	131	81	1203	393
<i>so and so</i>	113	317	219	363
<i>whatsis</i>	71	25	101	34
<i>whatsername</i>	49	13	3	12
<i>whatsisname</i>	13	9	7	8
<i>thingummy</i>	8	8	17	7

Table 3: Frequency survey based on all sources

4.6 Strategies in use

Emergent from the data were the following strategies in using placeholders:

- a) to refer to objects whose name escapes the speaker, or is on the tip of speaker's tongue but beyond linguistic reach;
- b) to keep listeners informed about our word retrieval problems;
- c) to deliberately withhold information; and

- d) to minimize risk and avoid interpersonal trouble (a self-protective behaviour, thread-avoidance strategies).

4.7 Reasons for using PHs

Emergent from the contextualized data were the following two reasons for using placeholders: *therapeutic* and *diplomatic*. Both are used with Goffmann's (1956) notion of face-work and the strategies *not to lose one's own face* and *not to threaten the face of the other*. The following samples will illustrate a couple of variations within therapeutic and diplomatic reasons of their occurrence.

4.7.1 Therapeutic use

Self-repair: the speaker identifies the delayed referent by him/herself

- (14) *Yeah. And somebody said, you know what, so and so deserves it, Collin Firth deserves it instead of you...* (BYU-COCA 2015 SPOK NBC Today Show)

Approximation of the missing referent: the speaker approximates the missing referent but it is the hearer who brings the word-formulation problems to a successful end

- (15) ...you can use er (sneeze) whatsit? It's not the dipstick it's the whatsit? – Joystick. *The joystick*. (laugh)

4.7.2 Diplomatic use

The following three strategies initiated the use of PHs:

- a) to deliberately withhold information;
- b) to minimize risk and avoid interpersonal trouble (cf. facework);
- c) to indicate shared social spaces (in-groupness; cf. *So-and-so*) and exclude those who are not part of the grouping.

Compare the following samples:

- (16) *Nobody is allowed to write about that. So-and-so will not be reprinted until we say so.* (BYU-BNC G1A W_fict_prose)
- (17) *What rabbi? Not that syrupy old Whatsisname you told me about at the beginning...* (BYU-BNC HGN W_fict_prose)

4.8 Relating the strategies to the most frequent PMs

Thingy is preferably used for the therapeutic, face protecting strategy, *so and so* – for a diplomatic withdrawal of the referent, and *whatsit* (as the most flexible of the three) is used

to manifest both the above mentioned strategies and it is the contextual setting that decides (cf. the above mentioned search for contextual clues).

Taking into view the results of the BNC and COCA corpora, we can state that there is an agreement in their ranking among the first three, but the positions in the frequency slots vary, cf. BrE – *thingy* (131), *so and so* (113), *whatsit* (71); AmE – *so and so* (317), *thingy* (81), *whatsit* (25).

We can tentatively state that the prevailing *thingy* in BrE signals the priority of a therapeutic use, while the prevailing *so-and-so* in AmE prompts the priority of a diplomatic use, with *whatsit* flexible to serve both the functions.

4.9 Summing up

With the existence of authentic language corpora, we are not at liberty to sweep evidence under the carpet. Vague language makes an important contribution to naturalness and the informal tenor of our everyday interactions. Placeholders, as one of the manifestations of vagueness are not deplorable deviations from precision and clarity: they can carry more contextual implications than precise expressions. As Channell (1994: 3) put it, “...vagueness in language is neither all ‘bad’ nor all ‘good’. What matters is that vague language is used appropriately”.

5 Conclusion

Language is constantly in a state of flux and must be continually negotiated. The multifacetedness and whimsical nature of discourse are properties that have to be taken into consideration in trying to understand the nuances of discourse in general and types of discourse in particular. In trying to isolate selected facets of discourse for research purposes we have to keep in mind that they operate in a complex interplay. The alchemy of discourse is partly predictable and partly innovative. Echoing Leech (personal communication) we can conclude by stating that discourse is a result of a linguistic juggling act in which we have to keep several balls in the air simultaneously.

References

- Biber, D., Connor, U. and Upton, T. (2007) *Discourse on the Move: Using Corpus Analysis to Describe Discourse Structures*. Amsterdam: John Benjamins. <https://doi.org/10.1075/scl.28>
- Blum-Kulka, S., House, M. and Kasper, G. (1989) *Cross-cultural Pragmatics: Requests and Apologies*. Norwood, New Jersey: Ablex.
- Austin, P. (1962) *How to Do Things with Words*. Oxford: Oxford University Press.
- Crystal, D. (2001) *Language and the Internet*. Cambridge: Cambridge University Press.

- Channell, J. (1994) *Vague Language*. Oxford: Oxford University Press.
- Enkvist, E. N. (1994) 'Centre and periphery, delicacy and fuzz.' In: Čmejrková, S. and Štícha, F. (eds) *The Syntax of Sentence and Text. A Festschrift for František Daneš*. Amsterdam: John Benjamins. 43–58. <https://doi.org/10.1075/llsee.42.06enk>
- Fillmore, C. J. (1992) '“Corpus linguistics” vs. “computer-aided armchair linguistics”.' In: Svartvik, J. (ed.) *Directions in Corpus Linguistics: Proceedings from a 1991 Nobel Symposium on Corpus Linguistics*. Berlin: Mouton de Gruyter. 35–66.
- Fillmore, C. J. (1976) 'Frame semantics and the nature of language.' *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech* 280, 3–25. <https://doi.org/10.1111/j.1749-6632.1976.tb25467.x>
- Goffman, E. (1956) *The Presentation of Self in Everyday Life*. Edinburgh: University of Edinburgh.
- Grice, H. P. (1975) 'Logic in conversation.' In: Cole, P. and Morgan, J. L. (eds) *Syntax and Semantics*, Vol. 3: *Speech Acts*. New York: Academic Press. 41–58. https://doi.org/10.1163/9789004368811_003
- Halliday, M. A. K. (1978) *Language as Social Semiotic. The Social Interpretation of Language and Meaning*. London: Edward Arnolds (Publishers) Ltd.
- Halliday, M. A. K. (1979) 'Modes of meaning and modes of expression: Types of grammatical structure and their determination by different semantic functions.' In: Albertson, D. J., Carney, E. and Holdcroft, D. (eds) *Function and Context in Linguistic Analysis*. Cambridge: Cambridge University Press. 57–79.
- Halliday, M. A. K. (2003) 'On the “architecture” of human language.' In: Webster, J. J. (ed.) *On Language and Linguistics*. Vol. 3. *Collected Works of M. A. K. Halliday*. London and New York: Equinox. 15–16.
- Halliday, M. A. K. and Webster, J. (2009) *Continuum Companion to Systemic Functional Linguistics*. London and New York: Continuum International Publishing Group.
- Lakoff, R. (1982) 'Persuasive discourse and ordinary conversation, with examples from advertising.' In: Tannen, D. (ed.) *Analyzing Discourse: Text and Talk*. Washington D.C.: Georgetown University Press. 25–42.
- Leech, G. N. (1983) *Principles of Pragmatics*. London: Longman.
- Leech, G. N. (2014) *The Pragmatics of Politeness*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195341386.001.0001>
- Mathesius, V. (1911) 'O potenciálnosti jevů jazykových.' (On the potentiality of the phenomena of language). *Věstník Král. čes. společnosti nauk*. 1–24.
- Quirk, R. (1987) *Words at Work. Lectures on Textual Structure*. Harlow, Essex: Longman.
- Searle, J. R. (1969) *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139173438>
- Schegloff, E. F. (1982) 'Discourse as an interactional achievement: Some uses of ‘UH HUH’ and other things that come between sentences.' In: Tannen, D. (ed.) *Analyzing Discourse: Text and Talk*. Washington, D. C.: Georgetown University Press. 71–93.
- Svartvik, J. (ed.) (1990) *The London-Lund Corpus of Spoken English*. Lund: Lund University Press.
- Tannen, D. (1982) 'Introduction.' In: Tannen, D. (ed.) *Analyzing Discourse: Text and Talk*. Washington, DC: Georgetown University Press. ix.
- Tárnyiková, J. (2009) 'English borrowings in Czech: Health to our mouths?' *Brno Studies in English* 35 (2), 199–213.
- Tárnyiková, J. (2009) 'Vague reference to notional categories (English-Czech interface).' In: Tárnyiková, J. and Janebová M. (eds) *AUPO Anglica III. Linguistica*. Olomouc: Univerzita Palackého. 115–132.
- Tárnyiková, J. (2010) 'Bags of talent, a touch of panic, and a bit of luck: The case of non-numerical vague quantifiers.' *Linguistica Pragensia* XX (2), 71–85. <https://doi.org/10.2478/v10017-010-0005-3>
- Tárnyiková, J. (2012) 'Halliday’s interpersonal component reconsidered.' In: Hopkinson, C., Tomášková, R. and Zapletalová G. (eds) *The Interpersonal Language Function Across Genres and Discourse Domains*. Ostrava: Universitas Ostraviensis. 26–36.

- Tárnyiková, J. (2019) 'English placeholders as manifestations of vague language: Their role in social interaction.' *Brno Studies in English* 45(2), 201–216. <https://doi.org/10.5817/BSE2019-2-10>
- Tárnyiková, J. and Válková, S. (2000) 'Politeness across cultures: The case of apologies.' In: Klégr, A. and Čermák, J. (eds) *The Tongue Is an Eye. Studies Presented to Libuše Dušková*. Praha: Univerzita Karlova. 101–116.
- Válková, S. (2004) *Politeness as a Communicative Strategy and Language Manifestation. A Cross-cultural Perspective*. Olomouc: Univerzita Palackého.
- Verschueren, J. (1999) *Understanding Pragmatics*. London: Arnold.
- Zipf, G. K. (1949) *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Cambridge, Mass.: Addison - Wesley Pub. Co.

Corpora

- Czech National Corpus ORAL2008*. Prague: Institute of the Czech National Corpus. Available from <http://www.korpus.cz/>.
- Czech National Corpus – InterCorp*. Prague: Institute of the Czech National Corpus. Available from <http://www.korpus.cz/>.
- Davies, Mark (2004) *BYU-BNC*. (Based on the British National Corpus from Oxford University Press). Available from <http://corpus.byu.edu/bnc/>.
- Davies, Mark (2008) *The Corpus of Contemporary American English: 450 million words, 1990–present*. Available from <http://corpus.byu.edu/coca/>.
- Davies, Mark's BUY suite corpora: corpus.byu.edu: The Movie Corpus and The TV Corpus.

Jarmila Tárnyiková is Professor Emerita of English Linguistics in the Department of English and American Studies, Faculty of Arts, Palacký University Olomouc. Her main research areas embrace present-day English, cross-language analysis of sentence complexes (English-Czech interface), text and discourse analysis, pragmatics and the language of fiction. Her current research focuses on corpus-based analysis of vague language and the strategies underlying its use. Member of SLE and the Prague Linguistic Circle, former Associate editor and current member of the Advisory Board of *Linguistica Pragensia*.

Address: Jarmila Tárnyiková, Department of English and American Studies, Faculty of Arts, Palacký University Olomouc, Křížkovského 10, 771 80 Olomouc, Czech Republic. [e-mail: jarmila.tarnyikova@upol.cz]

REPETITION AS AN EFFECTIVE RHETORICAL DEVICE IN REMARKS AT MEDIA BRIEFINGS ON COVID-19

Zhuzhuna Gumbaridze

<https://doi.org/10.5817/CZ.MUNI.P280-0212-2022-2>

Abstract

The present paper aims at exploring the use of a linguistic means of repetition as a persuasion strategy to achieve a communicative intention in a particular kind of discourse: addresses and opening remarks on COVID-19 related issues delivered at WHO and NCDC media briefings. The main objective is to investigate pragmatic function of repetition as a rhetorical device by means of which the text is assigned intentionality and the address becomes persuasive and manipulative as a consequence. Drawing on the assumption that such addresses lack a conversational space in which interactants would equally participate, a speaker takes a tough stance to bring forward ad hoc issues by utilizing repetition tactically and pervasively. This serves as a contributing factor to strengthening credibility of a speaker's interpretation of the pandemic crisis and actions proposed. The study highlights that while attempting to persuade, threaten, frighten or deter the audience into sharing a particular opinion of the state of affairs and undertake a proposed action, the speaker utilizes a direct appeal to the audience with the aim of having a specific impact on their opinions or behavior. In such addresses repetition functions as the main linguistic device used for rhetorical purposes.

Keywords

repetition, persuasion, parallelism, stance, rhetoric

1 Introduction

Even though rhetoric in human interaction is a construct that has been a topic of interest in linguistics for decades, recent trends have laid less emphasis on the use of repetition as a rhetorical device to transmit a communicative intention in discourse in which the speaker is assigned a specific identity and role. The present study endeavors to analyze the function of repetition as a communicative strategy used by speakers to influence the views and beliefs of the target audience so as to guide them towards an intended interpretation of the discourse and acknowledgement of reality. The research issue is undertaken primarily from the perspectives of discourse analysis and is based on a corpus of speeches delivered by Tedros Adhanom Ghebreyesus (Director-General of the World Health Organization (WHO)) and addresses given by Paata Imnadze (Deputy Director-General of the National Center for Disease Control and Public Health of Georgia (NCDC)). In terms of content and rhetorical

style, these authoritative actors are acknowledged authors of the speeches analyzed in the paper. The study of Ghebreyesus' remarks has been carried out on the transcript retrieved from <https://www.who.int/director-general/speeches> whereas the analysis of Imnadze's addresses has been based on the manually prepared transcript of the video records retrieved from <https://www.ncdc.ge/#/home> and https://www.youtube.com/watch?v=-NT9H_UKgr8&t=1s.

The paper asserts that a communicative strategy of repetition assigns the text not only an informative value, but also evaluative and persuasive consistency. A rhetorical device of repetition, which is pervasively used in addresses and opening remarks at press briefings on COVID-19, serves to represent the speaker's knowledge, views and beliefs about the pandemic crisis and the actions proposed and imposes an intended interpretation of the novel Coronavirus reality on the audience. The present paper uses qualitative and comparative research methods and is based on the study of the corpus which comprises two opening remarks made by Tedros Adhanom Ghebreyesus and two addresses delivered by Paata Imnadze in the media setting in 2020. The study is carried out from a contrastive perspective aiming at highlighting differences in the strategic means used by these two speakers to persuade the audience to accept their evaluation of the state of affairs. More specifically, the study explores contribution of the linguistic device of repetition to adopting a tough stance on conveying the pandemic circumstances by the speakers in the selected speeches.

2 Repetition for rhetorical purposes

A growing number of studies have been attempting to investigate the phenomenon of repetition (Norrick 1987, Tannen 1987, 1989, Johnstone 1991, Murata 1995, Rieger 2003, Bazzanella 2011, Evstafiadi 2020, Jancylo 2020) for years. Although this linguistic means has been a focal point of continual scrutiny and interpretation so far, rhetorical purposes of repetition and its contribution to strategic discourse lack a thorough study. It should be noted that in more recent years, the researchers' attention has mostly been directed toward the study of pragmatic functions of repetition in political discourse. Yet, repetition as a rhetorical device for persuasion strategies in the genre of opening remarks and addresses has not been given considerable attention. This remains a relatively unexplored and promising area of linguistic studies. An issue of interest that the paper aims to raise here relates to defining the function of anaphoric and parallel patterns correlating with a speaker's beliefs and norms of conduct in public discourse genres such as speeches, addresses and opening remarks. Anaphora and

parallelism are used by a speaker as a communicative strategy in their quest for making the message credible and persuasive and enable a speaker to qualify an attitudinal commitment to the truth of a proposition or actuality of a fact as such. Csicic defines the term anaphora as “an ancient rhetorical term for certain figures of repetition: synonymy, parallelism, epanaphora, iteration, repetition, epiphora. In its wider, but today obsolete meaning, it is the same as parallelism – in its narrower meaning, as a lexical anaphora... Anaphora is the repetition of lexemes at the beginning of clauses or parts of clauses which indicate the speakers’ stances and the communicative goals they achieve” (Csicic 2013: 71–73).

To achieve a communicative goal, utilizing numerous persuasive tactics is a common aspect characteristic to public addresses. Persuasion is regarded “as a highly complex, interactive, intentional and genre-specific aspect of discourse in which participants negotiate for meaning and adapt their perception of coherence to that of others” (Dontcheva-Navratilova 2017: 28). By attempting to persuade the listener to share his/her standpoint, the speaker imposes his/her ideology on the addressee. The choice of linguistic devices contributing to the persuasive force of a speaker’s rhetoric depends on the way he or she intends to legitimize what is proposed in the interaction. “The linguistic and rhetorical devices associated with persuasion are deontic and epistemic modality, expressed by modal verbs, nouns, adverbials, and clausal forms, such as *I think* and *I believe*, the choice between averral and attribution, associated with the use of intertextuality and reported statements, and the choice between assigning the status of fact, assumption or hypothesis to statements or larger parts of discourse” (Dontcheva-Navratilova 2017: 53). It is assumed that among various rhetorical and linguistic means including modality markers, the role of repetition as a stylistic device for influencing the beliefs or behavior of others is remarkable. The study of this linguistic phenomenon reminds us of Perkins’ attitude towards modality. Perkins remarks that “doing research on modality is very similar to trying to move in an overcrowded room without treading on anyone else’s feet... It is, after all, through treading on the feet of one’s predecessors (and sometimes through standing on their shoulders) that one discovers new doors to open” (Perkins 1983: 4). Similarly, the paper claims that despite vastness of literature on persuasion strategies, there is still much to focus on in regard to rhetorical repetition. The study will look briefly at the definition of this linguistic device and discuss what has been proposed so far.

The literature review shows that traditional understanding of repetition has been modified and the function of intensification of this linguistic means is being put forward. It is

claimed that intensification is a primary pragmatic function of repetition which evokes expressiveness and contributes to creating suspense and resolution. According to Bazzanella (2011: 249), “repetition, besides being a useful cognitive device (as a simplifying/clarifying device, a filler, and a support both for understanding and memorizing), an efficient text-building mechanism, and a widespread literacy and rhetorical device, is a powerful conversational and interactional resource”. Hsieh (2011: 163) argues that “pragmatically speaking, repetition, both self-repeats and other repeats, can be used to double up the illocutionary force, i.e., to do emphasis or to do persuasion, by means of repeating the linguistic form”. Tannen (1987: 97) describes repetition as “a resource by which conversationalists together create a discourse, a relationship, and a world. It is the central linguistic meaning making strategy, a limitless resource for individual creativity and interpersonal involvement”. According to Johnstone (1993: 27), “in the context of English stylistics, repetition is traditionally seen as a figure of speech. In other words, repetition is one way in which syntax can be made to deviate from the norm, so as to create a marked structure with a special rhetorical effect, or simply for ornamentation”.

An important insight to be applied to our understanding of repetition lies in defining repetition as a stylistic device the effect of which resides in highlighting the values, enhancing the proposition and persuading the addressees to align with the stance posited in the utterance. Then, why not treat this rhetorical device in terms of idea-making strategy in public appeals and addresses since this device is meant to guide the addressees’ interpretative processes of the state of affairs in accordance with the speaker’s intention? We claim that efficacy of repetition as a persuasion strategy to make others accept the speaker’s stance is really high.

3 Repetition as a persuasion strategy in addresses and opening remarks on COVID-19

Repetition is supposed to be a salient idea-making strategy that evidences a speaker’s more or less subjective evaluation of a proposition. It enhances persuasive rhetoric and helps a speaker realize a communicative intention artfully through asserting his proposition with varying degrees of certainty about the state of affairs. To see to what extent repetition is pervasive, powerful and functional in public addresses, a number of excerpts from speeches, addresses and opening remarks delivered by Tedros Adhanom Ghebreyesus and Paata Imnadze will be analysed.

- (1) *When the whole world got ill, and when we all became hostages of this virus, we have seen the centrality of it. **It affected the economy. It affected our social fabric. It affected our politics. It affected everything**.... Then, the other aspect of preparedness is, **I think this virus has exposed our societies, exposed us, exposed the problems that we have, the level of inequity, the level of poverty, the level of exclusion.** ...Then there was a very personal question about what **I think** about 2020. 2020 for all of us has been **a very difficult year, a very tragic one.** And for me personally, it's not just COVID alone. **COVID has been tragic, and is still tragic** for all of us. But personally, in addition to COVID, 2020 has been **very difficult** for me because **my country is in trouble. My country, Ethiopia, is in trouble.** And the war, the devastating war that's happening, is actually **in my home country, in my home region, Tigray, northern part of Ethiopia.** And **I** have many relatives there, including my younger brother, and **I** don't know where they are. **I** haven't communicated with them because communication is not there. As if COVID is not enough, **I** have that personal pain also. **I worry about my country.** Of course, my younger brother, or relatives, are part of the society. **I worry about the whole country. I cannot worry about my younger brother or relatives alone,** because the situation is worsening. (Ghebreyesus, 28 Dec. 2020) (https://www.who.int/docs/default-source/coronaviruse/transcripts/20-12-28-covid-19-vpc.pdf?sfvrsn=fc08ded0_1)*

By anaphoric repetition of *it affected* in successive sentences, Ghebreyesus presents factual information and attempts to appeal to addressees' ability to reason. The anaphoric repetition of the lexemes *country* and *tragic* points to the fact that COVID-19 has brought a great deal of suffering to people everywhere. The economic and social disruption caused by novel coronavirus is devastating. The pandemic has imposed inequity and poverty on everyone. The state of affairs has become an unprecedented challenge globally. Then, he brings examples of his personal connection to the subject matter and persuades the audience by appealing to their emotions. The repeated use of the personal pronoun *I/my* (*I worry, I cannot worry*) shows personal involvement and concerns about the issue. The linguistic strategies used by the speaker comprise emotive lexis and pronominal inclusive language. In the aforesaid fragment, the pathos is a predominant mode of persuasion in order to gain the audience's sympathy.

- (2) *So, carrying all this could be **tough, and has been tough.** Then **I see also hope.** Because in difficult situations like this, we can also find ways to build a better world and hope to improve situations and, as we all say, build back even better. During this time, by the way, two months ago exact, I became a grandparent, a grandfather. Although I worry, considering the two difficult situations that are happening, about my granddaughter, but at the same time, looking at her, **I see also hope.** And that's why in my speech, I said, if the world works together, **there is a solution. There is a solution globally, and there is a solution for any country.** So, **I am hopeful.** And whatever the situation is, I believe it can change. And I call on the global community **to choose peace, to choose solidarity, to choose caring for one another, caring for each other, and to choose without caring for one another, or each other,** that our world could be in trouble. And we have seen that. So that's what **I hope for.** And this is **the last session we have, the last presser** of this year. **And I hope** the problems and difficulties of 2020 will be behind us and we will have a better 2021. Even with the tragic events, I have said **I am hopeful** and we can do it together. **I have faith in each other. I have faith in us. I have faith in the world. And I hope** we will change for the better. I thank you.*

(Ghebreyesus, 28 Dec. 2020) (https://www.who.int/docs/default-source/coronaviruse/transcripts/20-12-28-covid-19-vpc.pdf?sfvrsn=fc08ded0_1)

The transcript exhibits pervasive use of repetition of various lexico-syntactic patterns. We assume that repetitive use is intended to give the text a rhetorical power and have a strong impact on the addressees. The text is obviously impressive, emotional, and convincing. Reduplication of either exact or similar syntactical patterns and lexical items in successive sentences indicates that Ghebreyesus wants the addressees to focus on ad hoc ideas. By repeating the words and using parallel structures, Ghebreyesus constitutes, asserts and reinforces significance of solidarity against the pandemic. Parallelisms (re-occurrence of syntactical and lexical similarities inside consecutive sentences) and anaphora (a rhetorical device involving the repetition of a word or group of words in successive phrases) are intentionally disseminated through the text by the speaker since he believes that these lively and vigorous linguistic devices will leave a strong effect on the audience. The rhetorical objective behind repetition in the last part of the excerpt lies in highlighting the sense of solidarity for the better future. The anaphoric repetition of lexemes aims at raising public awareness about the fight against the virus through maintaining unity. Ghebreyesus calls on the addressees to believe in their strengths, set objectives and strive for victory over the pandemic. He believes that combatting the common enemy of mankind is achievable through cooperation, support, sympathy and global solidarity.

By activating the context of hope, the speaker expresses an optimistic position towards the future. Repetitive use of anaphoric structure *I have faith in each other, I have faith in us, I have faith in the world* asserts his trust in a larger audience including mankind as a whole for the better future.

Let us consider another excerpt from opening remarks delivered by Tedros Ghebreyesus.

- (3) *There is no sacred in who because it's about lives and **life is very precious, even a single life is very precious. When we get information in DeBlasio, we get information in numbers: this many people died or this many people had have been infected. We don't see the figures, we don't see the numbers, we see the individuals behind the numbers, we see the individuals behind the figures. Of course, we may have averages on this, on that, but we don't want to see the averages. We want to see than numbers anyone dying if the daughter or a son of somebody, anyone dying is the father, mother, grandfather, grandmother of somebody. Its life and that's why we prefer to see the individuals, the faces, the people. They're not numbers, they're not averages. There are people, they're individuals and that's why we don't have secrets and as soon as we get information, we pass it because we want to save lives, even if it's one life, even one life matters, one life, you can't bring it back once it's gone and that's why we have been urging countries, please this virus is dangerous, this virus is public enemy number one, this virus is new and which has a behavior of serious contagion***

like flu it's very contagious, like flu and at the same time it's very killing like SARS and mercy it has a very dangerous combination and this is happening in hundred years for the first time again like the 1918 flu that killed up to 100 million people but now we have technology. We can prevent that disaster, we can prevent that kind of crisis, we can prevent it. We are not in the same situation, we should not be afraid, we should have the confidence that we're in a different situation and fight it back. If there is national unity and if there is global solidarity, if we take this as a common enemy for Humanity and give our best of course understanding that this is a new virus and dangerous virus we can win the fight, we can but please let's consider those who are dying as individuals. They are not numbers or figures. Even one life is precious. You know where I come from. I know war. I know poverty. I know disease. I know how people have suffered in all conditions. I know how people are killed because of poverty. I know how people who could have been saved are dying because of this. I know a tragedy. I know. That's why I see people. I see faces. I see the mother or father of somebody. I see the daughters and sons, sons of somebody, I see the tragedy that can behold a family. I see that. I see. Then, I know I'm telling you from a first-hand experience. I know how losing a brother means from childhood. I can tell you all the tragedies. I have seen, that's why in my last prayer I said let's not play with fire, this is a tragedy which is already affecting many families. Many are losing their loved ones. So, we don't hide information because I know what poverty means I know what war means I know what killer disease means I know what is behind all these problems and have been warning, we have been warning from day one. This is a devil that everybody should fight. Then the solution we're proposing we need national unity, strong national unity. Everybody fighting this virus taking care of their citizens, taking care of real people. We need global solidarity that cemented on genuine national unity without the - without national unity and global solidarity, trust us, the worst is yet ahead of us. Let's prevent this tragedy...

(Ghebreyesus, 20 Apr. 2020)

(https://www.who.int/docs/default-source/coronaviruse/transcripts/who-audio-emergencies-coronavirus-press-conference-20apr2020.pdf?sfvrsn=b5656a70_2

<https://www.youtube.com/watch?v=obuN4gPO97o&t=0s>)

The given example clearly demonstrates repetitive use of inclusive *we* that aims at establishing a common ground and building a rapport between a speaker and addressees. A speaker purposefully integrates himself into the audience in order to form unified relationship with the addressees. This “unification strategy is one of those prominent techniques utilized by politicians and rhetoricians in political text/talk; the planned deployment of *we*-groups (*we*, *our* and *us*) serves to create an affiliated relation between orators” (Derakhshani et al. 2021: 129). Noteworthy, by shifting from *we* to *I* in a successive syntactic string, Ghebreyesus places an emphatic stress on bad experiences of war, poverty and disease he has suffered and brings this contrast into light. The repetitive use of evaluative structures such as *I know* and *I see* shows a strong personal involvement in the issue and state of affairs. He reinforces the idea and importance of global solidarity and national unity in the struggle against the virus. Linguistically, at the lexico-semantic level, Ghebreyesus utilizes synonymous lexis in paradigmatic relations in order to activate an idea-making strategy. As regards the syntactic level, the parallel structures in the given examples are achieved by mood and textual patterns. Involvement with the audience is enhanced by occurrences of the

personal pronoun *I* with nouns referring hardship (war, poverty, disease, suffer), emphasizing the fact that Ghebreyesus is not distanced from the audience. In order to claim common ground, the speaker discloses himself and provides the examples from his personal experiences of difficult times. He guides addressees through the argumentation presented in the speech through emphasizing that what is being uttered is not just the speaker's beliefs and views, but the stance that is solidified by shared experiences, common ground and identical values. This demonstrates a high level of commitment to the values represented in the speech and enhances a persuasive force of the above discourse. Tedros is confident, committed and emotionally sensitive in his utterances. The rhetorical implication of his speech is that global solidarity is the most powerful weapon against the virus. He uses anaphora and parallelism as a persuasion strategy to preach about the role of national unity and solidarity in the state of affairs of pandemic reality. By a linear thematic progression of issues, the speaker guides the audience gradually to the newly-gained awareness of the collective responsibility of mankind against the virus. Thus, this linguistic configuration strengthens and empowers both the message and the moral of the entire speech.

As Dontcheva-Navratilova (2017) points out, on the textual plane, first person pronouns such as *I* and *we* serve as meta-discourse markers and aim at leading the addressee to the argumentation by activating relevant intertextual connections that is based on the participants' background knowledge. She argues that while the first person singular pronoun *I* is used as the instrument of self-disclosure "claiming common ground, narrative of belonging, humor, claiming authority and direct appeal to the audience and addressee, self-reference by the first person plural pronoun *we* is related primarily to the strategies of unification of in-group as opposed to out-group perceived as victim or threat/enemy, narrative of achievements and casting the present as a natural extension of the past" (ibid.: 125). More specifically, utilizing sensitive pronominal forms including inclusive *we* and repetitive patterns of emotively loaded key words in adjacent sentences signifies pathos of the utterance and intends to influence, modify, solidify or coordinate the addressees' interpretation of beliefs, attitudes and behavior.

Now, let us consider the way a tactical use of repetition helps Paata Imnadze to boost legitimation of his proposition and make people believe in the ideas and actions he proposes. To illustrate, the fragment from Imnadze's address is provided:

(4) **დღეს**, მთელი საქართველო რისკის ჯგუფია. **დღეს ჩვენ** ამის დრო აღარ გვაქვს. **დღეს ჩვენ** ვართ იმის საშიშროების წინაშე, რომ ჩვენთან განმეორდება იტალია. **დღეს**, ხვალ ჩემს თანამშრომელს ასაფლავებენ - ზურა მირიანაშვილს. კაცი, რომელსაც ოთხი სტენდი ჰქონდა გულში ჩადგმული. და ყოველდღე, დღეს და ღამეს ათენებდა და მუშაობდა. ვერ გავაჩერეთ, იმიტომ რომ იცოდა თავის სამშობლოს ემსახურებოდა, თავის ქვეყანას ემსახურებოდა.

ზურას ხსოვნის გულისათვის, **ხალხნო**, ნუ გავაკეთებთ იმას, რომ ის, რაც გავაკეთეთ დღემდე და გავაჩერეთ, მაქსიმალურად გავაჩერეთ, წყალში ჩაყაროთ ეს ყველაფერი...**გავიგოთ, გავითავისოთ**, რომ დისტანცირება ოჯახშიც კი არის საჭირო არათუ გარეთ. ესე, ჩვენ ცოტა ხანში, სამ კვირაში, საშინელებას დავატრიალებთ, თუ ეს არ გავითავისეთ **სულ ყველამ**. კიდევ ერთხელ მოგიწოდებთ **სულ ყველას, დარჩით სახლში**, არ წახვიდეთ არსად, და **მოერიდეთ** ერთმანეთსაც კი. თუ ვინმე ოჯახიდან უნდა გავიდეს და დაბრუნდეს ხოლმე, იმასაც **მოერიდეთ!** ის იყოს იზოლირებული ოჯახის სხვა წევრებისაგან! ასე ვრცელდება ეს ვერაგი დაავადება.

გაიგეთ **ხალხნო! ჩვენ** სხვა გზა არ გვაქვს! თუ არ დავრჩებით სახლში **ჩვენ** ვერ გადავრჩებით! **ჩვენ** გვექნება უდიდესი მსხვერპლი, აი ეს ყველამ უნდა **გავიგოთ!** თუ საჭირო გახდა, ზომები კიდევ უფრო გამკაცრდება!

...და თქვენგან ჩვენ **მეტი არაფერი** არ გვინდა, **დარჩით სახლში, მეტი არაფერი, დარჩით სახლში**. ჩვენ გავაკეთებთ ჩვენ საქმეს, **მეტი არაფერი არ გვინდა, დარჩით სახლში!** თავის საქმეს მიხედოს! ადამიანის საქმე არის, რომ **დარჩეს სახლში, დარჩნენ სახლში** და ეგ არის მაგათი მოვალეობა!

(Imnadze, 15 Apr. 2020) (https://www.youtube.com/watch?v=Gd0v_jOVCXY)

Imnadze combines a narrative with grief about the loss of his colleague Zura Mririanashvili. He grieves over the loss of the colleague because it was not only the loss of

Today, the entire Georgia belongs to a risk group. Today, we have no time for this. Today, we are in danger of sharing the experience of Italy. Today, tomorrow, my colleague Zura Mririanashvili will be buried. The man who had four heart stents placed in the heart, the one who worked days and nights every day. We could not stop him because he knew that he served his homeland, he served his country. For the memory of Zura, folks, don't make that, what we have done and stopped so far and have stopped maximally, go in vain. Let's be aware, be aware that distancing is necessary not only outdoors but also at home.

This way, we will meet with a disaster in a little while, in three weeks unless we, let's be aware of that. Appealing to all of you once again, stay at home! Don't go anywhere and keep a distance even from each other! If a family member has to leave home and then come back, keep a distance even from him! He must be isolated from other members of the family! This is how this violent disease is spreading. Be aware of that, folks! We have no other way but stay at home. We will not survive if we do not stay at home! We will have the greatest loss! This is what we have to be aware of! If necessary, the measures will get stricter!

...We want nothing else from you except staying at home. Stay at home! Nothing else. Stay at home! We will do our job! We want nothing else. Stay at home! Everyone ought to mind their own business! A human's business is to stay at home! To stay at home. This is their duty!

a peer or a friend, but also a very committed fellow front-line fighter against the coronavirus in this crisis. When seeking the way to convey authoritativeness, Imnadze constructs argumentation by conveying judgments about personal attitudes, beliefs and mandatory actions. Unlike Ghebreyesus, he is categorically imperative and provides the rule of conduct *Stay at home!*, which is an unconditional and absolute order for all addressees in saving lives. The repetitive use of this command enacts the speaker's persistent stance upon the state of affairs and discourages any disagreement or question from the addressees' side. His remarks express high degree of commitment to the proposition he utters. He claims obedience and approval of the actions proposed. He builds up the authoritative image of himself and the institution he represents and guides the audience towards an intended interpretation of the discourse. Paata speaks in an institutional voice and expects the audience to agree with his views, follow his instructions and obey stay-at-home orders unconditionally. The main persuasion devices used in this fragment are vocatives, direct commands and the first person plural pronoun *we*. Self-reference by using *we* is instrumental in the construction of the institutional identity of the speaker as a head of the organization Mr. Imnadze represents. The inherently subjective emotions and attitudes to the state of affairs are conveyed by the expressive lexis with negative connotative meanings, personal intrusions, a categorical tone and pessimistic pathos.

(5) ველარ დავითვლით საქართველოში კუბოებს!
 აი ეს გვინდა? **დავრჩეთ სახლში!** ხალხნო, აბა ჩვენ რატომ ვლაპარაკობთ იმას, რომ ქვეყანაში არის ფართო შიდა გადაცემა დაწყებული? ეს რას ნიშნავს? იმას ნიშნავს, რომ დღეს ჩვენ გვერდით არიან ინფიცირებული ადამიანები, რომლებმაც არ იციან, რომ არიან ინფიცირებულები. ისინი გადამდები ხდებიან უკვე სიმპტომების დაწყებამდე ერთი დღით ადრე, ანუ შეიძლება იმ დღეს საერთოდ არაფერი არ ჰქონდეს. იმიტომ ვამბობთ ამდენი ხანია, რომ შიდა გადაცემა დაიწყო ამიტომაც ვამბობ, რატომა ეს ყველა ღონისძიება, რატომ ტარდება. იმიტომ, რომ შიდა გადაცემა ქვეყანაში მიდის.
 გავიგოთ ეს!! ჩვენ გარშემო არიან ადამიანები, ინფიცირებულები, რომლებმაც არ იციან, რომ არიან ინფიცირებულები. ისინი მოვლენ თქვენთან

*We will not be able to count coffins in Georgia! Do we want that? **Let's stay at home!** Folks, why are we claiming that internal **transmission has already started** in the country? **What does it mean? It means that there are the infected** around us, who do not know that they are **infected**. They get contagious a day earlier the symptoms begin. For this reason, on the very day, they might have no symptoms at all. That is why we have been stating for so long **that transmission has begun**. That is why we are stating...why are all these measures being taken? Because **the internal transmission is going on** in the country.
Let's realize that! Around us, there are those who are infected and they do not know they are infected. They will visit you and make all of you infected! They will go to church and make all of those being there infected! **Let's realize that!** If you want to survive to stay alive, **stay at home!** **There is no other way out! There is no other way!** **Stay at home if you want our nation to survive, if you want the church to survive!***

სტუმრად და თქვენ გაგხდით ყველას
ავად. ისინი მივლენ ეკლესიაში და გახდით
იქ ყველას ავად. გავიგოთ ეს! თუ გინდათ,
რომ ცოცხლები გადავრჩეთ, სახლში
დარჩით! მეტი გამოსავალი არ არის, მეტი
არაფერი არ არის! სახლში დარჩით თუ
გინდათ, რომ ჩვენი ერი გადარჩეს, თუ
გინდათ რომ ეკლესია გადარჩეს!

(Imnadze, 15 Apr. 2020)

The message in the speaker's statement assures recipients that they are obliged to perform recommended actions if they want to survive. The speaker urges the recipients of the message to understand that staying-at-home position will protect them from the COVID-related negative consequences. While positing his claim, the speaker instills fear in addressees as he believes that exposure to a fear-arousing appeal will lead individuals to desirable behaviors. Strong emotive lexis (*death, not survive, coffin, victim, loss*) that is harshly used in adjacent clauses reinforces fear-arousing appeals and urges the audience in acknowledging the reality as a matter of life and death.

It is argued that rhetorical repetition (*Around us, there are those who are infected and they do not know they are infected. They will visit you and make all of you infected! They will go to church and make all of those being there infected! Let's realize that! If you want to survive, stay at home! There is no other way out! There is no other way! Stay at home if you want our nation to survive, if you want the church to survive*) creates a strong image of the threat of sinister consequences and is persuasive in the sense that it gives the audience a reason for making a choice between life and death.

Imnadze delivers his speech in a quite emotional but rather aggressive and categorical way. He requires the audience to follow his demand and comply with his orders. Involvement with the audience is reinforced by occurrences of *we*. By intensive inclusion of the first-person pronoun *we*, the speaker strives to establish a direct contact with the audience. He uses repetition saliently as a persuasion strategy to build his rhetoric. The speech is delivered in an emotional way. In order to enhance credibility of his messages, Imnadze overtly associates Corona crisis with excessive deaths in Georgia and builds up a whole set of death-related vocabulary such as *coffin, mortality, death, loss, victim, graveyard, grief, mourning, hospital, bury*, etc. By using anaphora and parallelism abundantly, Imnadze tries to persuade the audience that in the case they fail to obey stay-at-home rules, the consequences will be disastrous, fatal and extremely painful for all. It can be suggested that the use of emotionally

intensified language, second person pronouns, modal verbs *ought to*, *will* and imperatives serves to support and strengthen a speaker's request or demand. Thus, Imnadze's speech is essentially imperatival and mitigated by moral appeals. Repetitions that are used in the excerpt can be treated as rhetorical repetitions. They do not only just represent old information or serve as cohesive devices for discourse, but also facilitate the addressees to rate the intensity of the utterance and make further judgments. This excerpt demonstrates how repetition of the clauses *there are those who are infected and they do not know they are infected. They will visit you and make all of you infected! They will go to church and make all of those being there infected and There is no other way out! There is no other way! Stay at home if you want our nation to survive, if you want the church to survive* can help elucidate threatening aspects in the utterance. This accounts for its primary purpose as attempts to intimidate an addressee. We might call such messages fear appeals as they attempt to arouse fear among the audience by emphasizing the potential danger, loss and harm that will befall individuals if they disregard the speaker's recommendations and requests.

Similarly, as was mentioned earlier, to enhance the dramatic effect of the pandemic, parallelism and repetition in successive sentences are abundantly used for rhetorical purposes by Ghebreyesus as well. The fragment below illustrates that parallelism and repetition are effective means of persuasion potential. The elements and meaning that are alike in form in adjacent sentences are taken as a signal that they are fulfilling the same role in the expression and have an equal value in the utterance with a moral appeal. They keep listeners on track and involve them in the thinking process.

- (6) *This is a moment for all of us to reflect on **the toll the pandemic has taken, the progress we have made, the lessons we have learned**, and what we need to do in the year ahead to end this pandemic... We learn **something new** every single day. **Sometimes good, sometimes challenging, sometimes down right surprising**, but all helpful... As the year closes and people around the world **raise a toast** to mark both the passing of the year and the dawn of a new one, let me **raise a toast** to science; may we share its results – especially the vaccines – fairly and equitably in the year ahead and **together end this pandemic!** Happy New Year and I thank you. **None of us can end a pandemic by ourselves but together we will end this pandemic... None of us can end the pandemic by ourselves. But together, we will end this pandemic.***

(Ghebreyesus, 28 Dec. 2020)

(<https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---28-december-2020>)

The starting sentence in the extract under examination includes elaborate syntactic parallelism *the toll the pandemic has taken, the progress we have made, the lessons we have learned*. The parallel structures of clauses are part of the substantiation of the argumentative

judgement. These are the clauses presenting the key notions on which the whole utterance is based. This fragment shows a mixture of description, narration, opinion and prediction. The subjective belief of the text producer is discharged in his words. Despite harsh reality, the speaker aspires for victory against the pandemic and believes that humanity will prevail over the coronavirus. The speaker's certainty about the bright side of the future is indicated by an impressive array of positive lexis (*progress, lessons learnt, new, good, challenging, surprising but helpful, raise a toast, share results, together*) and modals (*may, can, will*) at the surface level. It clearly makes sense semantically to interpret the text as claiming that victory against the pandemic is not an illusion. The speaker is confident in the ability to tackle the global challenge through solidarity and the rhetorical purpose behind repetition lies in emphasizing his chief concern: *we will end this pandemic*. The repetition *None of us can end a pandemic by ourselves but together we will end this pandemic... None of us can end the pandemic by ourselves. But together, we will end this pandemic* is intended to engage the audience in a critical examination of the significance of unity. This mechanism profiles the speaker's optimistic beliefs and serves to encourage the audience to see the bright side of alternative perspectives of the future.

The discussion confirms that the addresses made by Tedros Adhanom Ghebreyesus and Paata Imnadze make a frequent use of similar linguistic strategies: the exact repetition of words, phrases, clauses, sentences and parallel constructions. What should be pointed out here is that repetition serves as an effective persuasion strategy both to communicate ideas and bring the addresses in rhetorical discourse. Pervasiveness of repetition in adjacent positions is interpreted as an attention-drawing verbal behavior from a speaker's side. This helps the speaker to go over ad hoc concepts, reactivate the context and extend it for further interpretation. In all the selected addresses, the speakers sound ostensive and emotional. Both speakers use anaphoric repetition and parallelism pervasively as a persuasion strategy in their addresses, but they construct different lexis to enhance credibility and commitment to the proposition. Through the careful selection of the positive vocabulary Ghebreyesus' remarks are optimistic, rational and hope evoking, whereas Imnadze chooses threatening language and makes pessimistic and fear-arousing appeals. By repeating words, phrases, clauses and sentences Imnadze plays on the effect of repetition to highlight pre-designed assumption of state of affairs to impose his beliefs on the audience by creating a sense of despair, threat and ominousness.

The analysis of the excerpts from Imnadze's addresses has confirmed that the rhetorical device of repetition serves to represent harsh reality, create a pessimistic scenario of the state of affairs and envisage hopeless consequences which aim at imposing fear, disappointment, despair and threat on the audience. By repeating words, phrases and clauses Imnadze assigns the argument a persuasive force and represents his commitment toward the proposition, ranging from description of the state of affairs to confident prediction of ominous consequences. Yet, among Georgian speakers, the use of this rhetorical device differs with an individual style. What is particularly important here is that repetition in both corpora makes the utterances more expressive and adds extra effects to the focal segment of the appeal.

It should be emphasized that while attempting to persuade, threaten, frighten or deter the audience into sharing a particular opinion of the state of affairs and undertake a proposed action, the speaker utilizes a direct appeal to the audience with the aim of having a specific impact on their opinions or behavior. In such addresses repetition functions as one of the main linguistic devices used for rhetorical purposes. The speakers' subjective emotions and attitudes towards the pandemic situation are linguistically conveyed by direct commands, vocatives, expressive lexis and personal intrusions on the part of the speaker by means of self-reference personal pronouns. All these elements are pervasively used in repetition patterns and contribute to persuasive effect of the utterance.

4 Conclusion

The analysis demonstrates that repetition is pervasively used as a powerful persuasive strategy in addresses, opening remarks and speeches on COVID-19 at the media briefings. Anaphora and parallelism as the most commonly utilized forms of repetition, are powerful rhetorical devices in addresses and opening remarks to convince addressees. However, unlike in the English corpus, by encoding a text with pre-selected linguistic devices, Mr. Imnadze not only represents the pandemic state of affairs as facts as such, but also constructs a pessimistic scenario of the country's pandemic future. Therefore, Georgian addressees are exposed to fear-arousing appeals that represent pessimistic views and anticipated sinister consequences of the pandemic circumstances. At Georgian COVID-19 related media briefings, reporting policy is pragmatically conveyed through using a categorical tone and creating a sense of threat and ominousness whereas Mr. Ghebreyesus uses more positive discursive strategies to describe either the pandemic reality or anticipated risks. Unlike in English, in Georgian, the speaker intentionally constructs a particular set of the target negative

vocabulary attributed to Coronavirus consequences and by repetitive use of such lexis in successive syntactic strings, the speaker both adheres an instructive power to the speech in a particular context and attempts to intimidate the audience by asserting that dreadful misfortune and loss of lives will befall them. In Georgian, repetition attains a strong illocutionary force and performs the speaker's purposeful communicative intention to create a sense of fear, threat, despair and ominousness.

The research proves that in opening remarks and addresses on COVID-19, repetition is used as a strategic means to perform persuasive and manipulative role in the formation of the attitudes of the audience. It follows that in both corpora, repetition is the means by which a speaker expresses a high degree of commitment to the truth of the proposition he utters and wants an addressee to perceive the states of affairs the way the speaker represents. Rhetorical and linguistic choices serve to legitimate the speaker's ideological views and guide listeners towards text interpretation which suits his or her communicative intention. Thus, anaphoric repetition in adjacent clauses represents a rhetorical device that is used by speakers tactically with the intention to persuade the people that they have maturity, strength, power, public consciousness and social responsibility to end the pandemic. I argue that repetition as a linguistic means is extremely effective for persuasion strategy when there is a rhetorical purpose behind using it. In this case, repetition is not only contextually motivated but also addressee-oriented and serves to encourage, motivate, praise, blame, warn and intimidate the audience.

References

- Bazzanella, C. (2011) 'Redundancy, repetition, and intensity in discourse.' *Language Sciences* 33(2), 239–242. <https://doi.org/10.1016/j.langsci.2010.10.002>
- Cisic, A. (2013) 'On anaphoras in political discourse.' In: Smetonienė, I. and Petrenienė, O. (eds) *Contemporary Stylistics: Trends and Problems. Collection of Scientific Articles*. Vilnius: Edukologija. 69–78.
- Derakhshani, M., Qaiwer, S. N., Kazemian, B. and Mohammadian, S. (2021) 'Critical discourse analysis and rhetorical tropes in Donald Trump's first speech to the UN.' *Theory and Practice in Language Studies* 11(10), 1224–1236. <https://doi.org/10.17507/tpls.1110.10>
- Dontcheva-Navratilova, O. (2017) *Coherence in Political Speeches*. Brno: Masaryk University. <https://doi.org/10.5817/CZ.MUNI.M210-8859-2017>
- Evstafiadi, O. (2020) 'Syntactic and pragmatic markers of subjectively charged English complex conditional sentences.' In: Pavlova, A. (ed.) *European Proceedings of Social and Behavioral Sciences: Philological Readings* 83. Orenburg: Orenburg State University. 55–61. <https://doi.org/10.15405/epsbs.2020.04.02.7>
- Hsieh, F. (2011) 'Repetition in social interaction: A case study on Mandarin conversations.' *International Journal on Asian Language Processing* 19(4), 153–168.

- Janczylo, A. (2020) 'The role of repetitions in Barak Obama's speech and its Polish translation.' *Forum Filologiczne Ateneum* 8(1), 227–240. [https://doi.org/10.36575/2353-2912/1\(8\)2020.227](https://doi.org/10.36575/2353-2912/1(8)2020.227)
- Johnstone, B. (1991) *Repetition in Arabic Discourse. Pragmatics and Beyond New Series (Paradigms, Syntagms and the Ecology of Language)*. Amsterdam and Philadelphia: John Benjamins Publishing Company. <https://doi.org/10.1075/pbns.18>
- Murata, K. (1995) 'Repetitions: A cross-cultural study.' *World English* 14(3), 343–356. <https://doi.org/10.1111/j.1467-971X.1995.tb00078.x>
- Norricks, N. (1987) 'Functions of repetition in conversation'. *Text - Interdisciplinary Journal for the Study of Discourse* 7(3), 245–264. <https://doi.org/10.1515/text.1.1987.7.3.245>
- Perkins, M. R. (1983) *Modal Expressions in English*. London: Frances Pinter (Publishers).
- Rieger, C. L. (2003) 'Repetition as self-repair strategies in English and German conversations.' *Journal of Pragmatics* 35 (1), 47–69. [https://doi.org/10.1016/S0378-2166\(01\)00060-1](https://doi.org/10.1016/S0378-2166(01)00060-1)
- Tannen, D. (1987) 'Repetition in conversation as spontaneous formulaicity.' *Text - Interdisciplinary Journal for the Study of Discourse* 7(3), 215–243. <https://doi.org/10.1515/text.1.1987.7.3.215>
- Tannen, D. (1989) *Talking Voices: Repetition, Dialogue, and Imagery in Conversational Discourse*. Cambridge: Cambridge University Press.

Zhuzhuna Gumbaridze is Associate Professor at Batumi Shota Rustaveli State University, Faculty of Humanities, Department of European studies, Batumi, Georgia. Her main research interests lie in applied linguistics, discourse analysis, stylistics, pragmatics and ELT methodology. Her dissertation dealt with discourse pragmatics and its verbal and non-verbal dimensions in Georgian and English languages. She is involved in various projects within the university focusing on quality enhancement. She is the manager of the bachelor educational program of English Philology, the author of numerous articles, the teacher trainer and the head of the curriculum committee at the faculty of Humanities, BSU.

Address: Zhuzhuna Gumbaridze, Department of European Studies, Faculty of Humanities, Batumi Shota Rustaveli State University, Ninoshvili Str. 35, Batumi, Georgia 6010. [e-mail: Zhuzhuna.gumbaridze@bsu.edu.ge]

TENDENCIES OF THREAT TECHNIQUES UTILIZATION IN THE COVID-19 RELATED DISCOURSE IN GEORGIAN AND ENGLISH ONLINE NEWS REPORTS

Zeinab Gvarishvili and Nana Mazmishvili

<https://doi.org/10.5817/CZ.MUNI.P280-0212-2022-3>

Abstract

The objective of the present research is to study and analyze linguistic markers of threat utilized in English and Georgian online daily news (*The Guardian* and *The Georgian Times*) concerning the COVID-19 pandemic. The study provides synchronic analysis of specific tendencies of utilization of threat techniques in the two languages. The research methodology implies qualitative and quantitative data analyses as well as corpus-based research aiming at determination of peculiarities of threat utterance and its function in persuasive communicative act and revealing similarities and differences between the two languages by contrasting and comparative methodologies.

Keywords

COVID-19 online news reports, persuasion, threat, linguistic markers of threat

1 Introduction

The issue of dramatic changes due to the COVID-19 pandemic has attracted interest of many scholars. Thus, throughout the period of its existence starting from 2020, an impressive amount of research have been conducted with the objective to elucidate outcomes of the breakout and spread of the virus not only in the specific areas of everyday life (education, economy, politics), but also the change of linguistic behavior manifested in everyday discourse.

Despite the profound number of studies, there is still a room for further linguistic research pertaining to linguistic markers utilized in the COVID-19 related discourse. Therefore, the aim of the presented study is to determine the strategies to communicate threat (used in online news reports) as one of the means of audience persuasion in effectiveness of the applied measures to withstand and prevent further spread of the virus. Thus, the objective of the present research is to study the linguistic techniques utilized to communicate threat (both explicit and implicit) regarding the COVID-19 virus and its possible complications to the target community, in English and Georgian online daily news. The study provides

a synchronic analysis of specific tendencies of threat technique usage. Namely, the present study aims i. to detect the types and frequency of usage of threat linguistic techniques in English and Georgian online news with the aim to define general tendencies of their application; ii. to elucidate similarities and differences between the two languages on the basis of cumulative evidence.

Since the breakout of the COVID-19 pandemic, it has become one of the biggest challenges to the whole world. Being the worldwide threat, the pandemic has altered the way people live in terms of social behavior, economic policies and political priorities, and some of these changes may stay with us for a long time. Under such circumstances, authorities had to take responsibility to identify, assess and respond to the risks brought by this novel virus. Countries throughout the world have framed the public health emergency as a national security threat and declared “national emergencies” over the COVID-19.

The Copenhagen School of Security Studies (2020) has formulated the notion of securitization. Central to this process is the invocation of the concept of survival. According to Buzan et al. (1998: 21–22), “security is closely linked to survival and everything that threatens the existence of a referent object can be potentially declared as a security issue”. Traditionally, security was only invoked when it came to external military threats that endangered national survival. This is one of the reasons why war rhetoric is used as one of the means of securitizing the COVID-19. War rhetoric can be easily detected in mass media pandemic discourse. As Kozlova (2021) states, the pandemic is interpreted in terms of war and the world is seen as a war zone where people are fighting against the disease. Thus, the words associated with war (*battle, beat, buckle, combat, conquer, defeat, fight, struggle*) are largely used in the COVID-19 related discourse. The relationship between the actors of the situation (the pandemic, virus, humanity, society, business, economy, etc.) is considered as a battlefield in which one party becomes victimized (*victim of the pandemic*).

The COVID-19 pandemic has been declared as the national emergency. This process is directly related, firstly, to recognition of the severity of the threat and, secondly, to acceptance of the necessity to follow the policy measures (quarantines, mask-wearing, social distancing, national lockdowns and vaccination) introduced by the authorities. Thus, persuasion of the audience in rationality and effectiveness of the applied measures has become of pivotal importance.

One of the most affective and powerful ways to gain public confidence and influence their behavioral change is the use of rhetoric. According to Aristotle (1939), rhetoric is

defined as “the art of persuasion” or the art of using language in such a way as to impress an audience, to persuade them for or against a desired course of action. Rhetoric can be broken down further into “non-artistic”, which comes from pre-existing, external facts, and “artistic” means of persuasion, embracing three core elements: ethos (ethical proof), pathos (emotional proof), and logos (logical proof). All three components are complementary since they activate specific sides of a listener’s mind and emotions. Artistic means of persuasion has been widely used in the COVID-19 related discourse to convey information about credibility and reliability of the COVID-19 policies (ethos), instil strength and courage (pathos), defend government interventions by providing reasons to believe in the risks posed by the COVID-19 epidemic, and persuade the audience of the effectiveness of the suggested measures (logos).

2 Literature review

2.1 Persuasion

According to Perelman and Olbrechts-Tyteca (1969), persuasion is a skill that human beings use in order to make their partners perform certain actions or collaborate in various activities. This is done by modifying – through communication (arguments) – the other’s intentional attitudes. In fact, apart from physical coercion and the exploitation of stimulus-response mechanisms, the only way to make someone do something is to change his beliefs. In persuasion, the speaker presupposes that the recipient is not already performing or planning the required action. It can also be presupposed that the recipient has some barriers against required action, hence, he/she would not spontaneously intend to do so. Persuasion is then concerned with finding means to overcome these barriers by conveying the appropriate beliefs to the interlocutor.

As it is defined by Simons (1976: 21), persuasion is a “human communication that is designed to influence others by modifying their beliefs, values, or attitudes”. O’Keefe (1990) argues that there are requirements for the sender, the means of persuasion, and the recipient to consider something persuasive. First, persuasion involves a goal and the intent to achieve the goal by the message sender. Second, communication is the means to achieve the goal. Third, the message recipient must have free will. Accordingly, persuasion is not accidental, nor is it coercive. It is inherently communicational.

Being one of the most powerful ways of rhetoric, the verbal act of persuasion is widely discussed in scientific fields of linguistics, psychology, behavioral science, etc. Scientific literature provides findings concerning four main theories that make a baseline of

persuasion technique: *social judgment theory, elaboration likelihood model, cognitive dissonance theory and narrative paradigm* (Festinger 1957, Sherif & Hovland 1961, Simons 1976, Petty & Cacioppo 1986, Fisher 1987, O’Keefe 1990, Cialdini 1993).

Social judgment theory and elaboration likelihood model consider the importance of being aware of the audience before creating a message. Social judgment theory suggests that successful persuasion can be guaranteed only if there are proper preliminary attitudes that allow the persuasion to occur “at the edge of his or her latitude of acceptance”. Knowing the audience is the key to successful persuasion. Based on the above-mentioned theory, acceptance of the message by the audience is directly connected with the ability to process the message. Cognitive dissonance theory argues that persuasion is a “postreactive response” in case the recipient struggles between the beliefs and the actions. As a rule, people are not in favor of incoherence between their beliefs and behaviors; therefore, persuaders can use the existing dissonance for their benefit and propose a solution that would make the audience feel that the inconsistency between their beliefs and behaviors is decreasing; thus, this persuasion tactic operates successfully. The narrative paradigm is amongst the list of persuasion tactics that is fully based on creating emotional picture through storytelling. Moreover, narratives must meet certain retelling requirements to maintain influential power throughout the plot of the story.

2.2 The role of threat in persuasion

As Mongeau (2013) and Rogers (1983) state, threat appeals consist of two major components: *the hazard* and *action components*. The hazard component involves two minor parts that describe (a) the undesirable consequences of an external threat (severity), and (b) the likelihood that the individual will experience these consequences (susceptibility). The action component is composed of (a) instructions for an effective response to an external threat (response efficacy), and (b) material concerning an individual’s ability to execute that response (self-efficacy). The two major components – hazard and action – are typically presented in a problem-solution format. Individuals are first presented a danger, and then exposed to the means of addressing the danger (Beck & Frankel 1981, Witte 1992). Judgments of severity and susceptibility provoke a particular emotion – fear. Thus, threat is regarded as a message, while fear is an emotional response that follows perceived threat. According to Hovland et al. (1953: 66), factors that make successful persuasive appeal are:

“1) factors for successful arousal of emotional tension-fear, and 2) factors for successful rehearsal and reinforcement of the communicator’s reassuring recommendations”.

3 Persuasion strategies in the COVID-19 related discourse

Taking into account the above presented theories, thorough analyses of authentic resources revealed that out of four persuasive strategies, the two – *Narrative paradigm* and *Elaboration likelihood model* – are the most widely used in the COVID-19 related news reports. The former focuses on emotional response of the audience (provoking fear) through storytelling, while the latter aims at rising motivation, reassuring the readers and persuading them in their abilities in terms of withstanding the pandemic if the provided regulations are duly followed.

The persuasion technique utilized in the COVID-19 related news report discourse is represented as a four-stage verbal act:

1. **Introduction of the threat.** This stage comprises
 - a. demonstration of the severity of the threat.
 - b. description of the threat using vivid word pictures illustrating the horrible consequences of the threat (severity).
2. **Personalization of the threat.** The stage comprises:
 - a. concentration on the risk of the threat to the audience.
 - b. placing emphasis on the similarities between victims of the threat and the target audience (susceptibility).
3. **Verbal demonstration of the effectiveness of the solution.** This stage covers
 - a. providing the solution that does not outweigh the importance of the threat.
 - b. providing accurate data about the audience who have already used the method (response efficacy).
4. **Insuring the audience in their capability to implement the solution** (self-efficacy). If the audience are not sure that the solution will work or that they cannot implement the solution, they will feel it is futile to try to control the risk. Thus, they will focus on controlling their fear rather than threat.

In order to provide these hypotheses with appropriate arguments, as well as to present an authentic comparative analysis of the English and Georgian online news reports concerning the use of linguistic markers of threat, the study aims at a thorough analysis of English and Georgian COVID-19 related online news reports.

4 Data and methodology

The research is based on the corpus-analysis study with the objective to identify the frequency of application of threat markers utilized via different persuasion techniques in the Georgian and English online news on the COVID-19 pandemic. For the comparative investigation online Georgian and English news articles on the COVID-19 pandemic were selected and analyzed through utilization of corpus study tools. The first stage was selection of online news on the COVID-19 pandemic in both languages. To guarantee the reliability and validity of the data obtained through the research, the online news items were of the similar content as well as objective. Moreover, the texts of a similar size (number of words used in each source) were chosen for the accuracy of the data for one month, namely December 2021.

The next stage envisaged identification of rhetorical means (metaphor, epithet, detailed or metaphorical description of the situations in the hospitals, mentioning and citing authorities – both national and international-providing statistics, delivering the message on privileges for the vaccinated community, utilization of conditional sentences, other implicit ways aimed at threatening and persuasion) in Georgian and English online news.

The next stage involved tagging of the identified threat markers and analyzing the tagged texts through the software *Orange*. Corpus-based analysis enabled us to identify i. the exact number of application of each threat marker in the selected empirical materials; ii. an *average power* (AP) of the marker – the number of words used within the tag divided by the number of words in the whole text selected for the analysis; iii. *density* of the applied tag (DAT) based on calculation of an average (AP divided by the overall number of the texts analyzed).

The results of the research reveal the types of threat markers used in both languages as well as the frequency of application of each threat marker within one language.

5 Discussion

The results of the analysis of English and Georgian online news reports revealed the utilization of a variety of linguistic techniques targeted at communicating threat to the recipient. Interpretation of corpus-based data helped to identify tendency in application of each type of linguistic marker from the most to the least dominant linguistic techniques both in English and Georgian online news reports.

As mentioned above, the analysis of the linguistic techniques utilized to communicate persuasion enabled us to have a comprehensive scenario of application of the most powerful persuasion techniques in the English as well as Georgian online news reports on the COVID-19 with the aim to see the pattern of strategic application of language in the two different cultural environments. The table presented at the end of the analysis of the types of persuasion techniques in two languages depicts the similarities and differences in the types of persuasion techniques and the frequency of application of each type of widely applied persuasive means in each language.

5.1 Threat markers: Average power and density in the English online news reports

The results of the data analysis revealed that English online news on the COVID-19 mostly utilize two persuasion strategies – *Narrative paradigm* and *Elaboration likelihood model*, which, as mentioned above, allow the audience to process logically the delivered message and react in the way the persuader wants them to. This is done through various strategies including statistics, conditional sentences, citing authorities and utilizing expressive means such as epithet and metaphor when describing the virus itself and its horrendous effect on people's health.

The analysis of the English online news demonstrated preference to the utilization of statistics as one of the effective ways of persuasion used to deliver the message about the negative influence of the virus on health. In most cases, the data are presented within the sentence or a paragraph in the body part of the online news report (Example 1):

- (1) *The COVID-19 hospitalizations jumped 88% in the past month, according to the Michigan Health and Hospital Association; "Since January, we've had about 289 deaths; 75% are unvaccinated people" Dover said.*
(<https://edition.cnn.com/2021/12/09/us/hospital-covid-19-deaths-michigan/index.html>)

One of the most frequently applied persuasion strategies considers provision of comparative scenarios for vaccinated and unvaccinated people with a clear message on privileges for the vaccinated community (Example 2):

- (2) *This week, data presented by a CDC vaccine adviser showed a hospitalization rate 16 times greater in the unvaccinated population than in those vaccinated; "Since January, we've had about 289 deaths; 75% are unvaccinated people," Dover said. "And the very few (vaccinated people) who passed away all were more than 6 months out from their shot. So we've not had a single person who has had a booster shot die from Covid."*
(<https://edition.cnn.com/2021/12/09/us/hospital-covid-19-deaths-michigan/index.html>)

The next widely applied strategy to persuade the audience in the English online news on the COVID-19 threats is metaphor, which is one of the types of techniques utilized under the narrative paradigm persuasion methods. The data obtained through the analysis of the empirical material provide evidence that proves the power of metaphor in communicating the message targeted at an emotional impact on the recipient. In respect of metaphors in the English online news it must be noted that majority of contexts dealing with metaphors expose the recipient to the harmful and severe nature of the virus. With this purpose, heart-touching descriptions of the situations created in the hospitals are used in the English online news. In these descriptions the application of metaphors, which assign human qualities to nonhuman things, creates threatening effect, thus persuades the reader to follow the rules. Reading passages about how a personified virus *strolls from a ward to a ward, takes lives* of people regardless their age, position, knowledge, etc. truly makes any recipient become overwhelmed by negative emotions, and thus leads to acceptance of the message of the persuader (Example 3):

- (3) *“So the virus is like a dangerous driver fleeing the scene - **the virus has moved on to the next victim long before we either recover or die;**” in stark terms, “**the virus doesn’t care**” if you die, says Prof Lehner, “**this is a hit and run virus.**”*
 (<https://www.bbc.com/news/health-54648684>)

The bold parts of the above given sentences are ontological metaphors that animate the virus which “behaves” as an invisible dangerous enemy the fear towards “which” motivates the recipient to act in the way desirable for the addresser.

One of the effective ways to make the message convincing is mentioning authorities as the source of information; this can vary from mentioning an authorized person responsible for provision of news on the current status of the virus to the local bodies, specializing in the concrete directions as well as international organizations dealing with world health issues in general. The more famous the institution is, the more trust the message supposedly gains in the recipient. In regard of citing authorized people or institutions, it is worth mentioning that there are two ways to name them: a. direct citation and b. indirect citation. For more accuracy it has to be clarified that both types of citations (direct and indirect) are applied in the case of authorized figures, while to show data of negative tendencies in spread of the virus indirect citation of the authorized organizations is applied (Examples 4 and 5):

- (4) ***Walensky said** that while people who are fully vaccinated can travel with precautions, current transmission rates mean they too need to take the COVID-19 risk into consideration when deciding whether to travel.* (indirect citation; citing an authorized person)
 (<https://edition.cnn.com/2021/09/01/health/us-coronavirus-wednesday/index.html>)

- (5) *“But this week Michigan had more patients hospitalized for the COVID-19 than ever before. The COVID-19 hospitalizations jumped 88% in the past month, according to the Michigan Health and Hospital Association.”* (indirect citation; presenting data from the authorized institution)
(<https://edition.cnn.com/2021/12/09/us/hospital-covid-19-deaths-michigan/index.html>)

When talking about citations, as one of the effective strategies to persuade the audience, it must be noted that citing hospital staff is one of the prominent techniques, and this is not surprising as “telling stories by witnesses” is considered to be a very powerful tool as ensured by the Narrative Paradigm. Moreover, direct citation of hospital staff conveys a message that mainly deals with metaphorical presentation of the hospital cases resulting in strong emotional response of the recipient. The below given examples can prove the influential nature of sad stories of individual persons affected by the virus (Example 6):

- (6) *“We’re seeing a lot of young people. And I think that is a bit challenging.” Said Sefton, a 20-year nursing veteran. She recalls helping the family of a young adult say goodbye to their loved one; “Katie Seftom said she’s now seeing younger patients die from the COVID-19”; “It was an awful night,” she said. “That was one of the days I went home and just cried.”*
(<https://saudigazette.com.sa/article/614568>)

As the research results show, conditional sentences follow the technique of telling stories or describing the situations in hospitals with the level of frequency of their application in the English online news reports (Example 7):

- (7) *“If you are unvaccinated, we would recommend not travelling,” CDC Director Walensky said; “That terrifying forecast could be prevented if more Americans get vaccinated,” Dr. Anthony Fauci, Director of the National Institute of Allergy and Infectious Diseases told CNN Sunday.*
(<https://edition.cnn.com/2021/09/01/world/coronavirus-newsletter-intl-1-09-21-duplicate-2/index.html>)

The analysis of the conditional sentences used to communicate threat revealed *if* as a prevailing marker as part of the persuasive strategy.

The results of the analysis of the English online news on the COVID-19 reveal that another stylistic device communicating threat is epithet; however, this is the least frequently applied strategy in the analyzed materials. For example, *killing virus*, *Covid-free territories* are the epithets that show the qualities associated with the virus.

5.2 Threat markers: Average power and density in the Georgian online news reports

The analysis of the Georgian online news on the COVID-19 revealed that Georgian texts most frequently mention authorized persons and institutions or organizations responsible for management of the virus spread and possible consequences on the country level.

Moreover, besides mentioning the names of local authorities from the Ministry of Health or other governing bodies, Georgian online news name international organizations as one of the powerful means to gain trust and ensure more reliability of the data (Examples 8 and 9):

- (8) *აშშ-ის ალერგიისა და ინფექციური დაავადებების ეროვნული ცენტრის ხელმძღვანელი ენტონი ფაუჩი აცხადებს, რომ ვირუსის დელტა ვარიანტი ძალიან გადამდებია და ამ ვითარებაში პრევენციის საუკეთესო გზა ვაქცინაციაა (Anthony Fauci, head of the US National Center for Allergy and Infectious Diseases, says that the delta variant of the virus is highly contagious and the best way to prevent it in this situation is vaccination.)*
(<https://www.radiotavisupleba.ge/a/31375013.html>)
- (9) *ძალიან ვთხოვე ჯანდაცვის ეროვნული ცენტრის ხელმძღვანელს, ბატონ ამირანს გარდაცვლილთა რაოდენობის გამოცხადებისას ისიც აღნიშნონ, რამდენია მათ შორის ვაქცინირებული (I asked the head of the National Health Center, Mr. Amiran, to announce the number of vaccinates, when announcing the number of dead people.)*
(<https://www.radiotavisupleba.ge/a/31375013.html>)

The research results confirm that it is effective to convince recipients both in harmfulness of the virus and the advantage of vaccination as long as they are given the examples of positive consequences of vaccination or exposed to bad statistics for unvaccinated community. With this purpose the Georgian online news reports often provide a contrastive picture with a clear message on different levels of harm the virus can do to those who got vaccinated and those unvaccinated (Example 10):

- (10) *როგორც ის განმარტავს, მათი კლინიკის 20 კაციან რეანიმაციაში დიდი ხანია აცრილი ადამიანი არ მოხვედრილა. რაც შეეხება გარდაცვალებას, გადელთა ამბობს, რომ მათთან აცრილი მხოლოდ ერთი გარდაიცვალა და ისიც მძიმე თანმხლები დაავადებებით*
(As he explains, they have not received any vaccinated people in the intensive care department with 20 beds of their clinic. As for deaths, Gadelia states that there was only one death case from those vaccinated who had serious accompanying diseases.)
(<https://www.ambebi.ge/article/269360-cvrilshviliani-ojazi-darcha-umamod-acrili-rom-qo/>)

In the Georgian online news, Narrative Paradigm appears as one of the widely applied persuasion strategy. The power of telling stories about the victims and survivors of the virus is obvious here too. In other words, personalized cases have more influence on the recipients than pure statistics that touches sensory modalities of a person less than exposure to dramatic scenes. Thus, telling stories of victims of the virus follows conditional sentences according to the level of frequency with which these means and techniques are utilized in the Georgian online news reports (Example 11):

- (11) *“...აპრილი არ არის ისეთი ცხელი თვე როგორც ივლისი, მაგრამ უჰაერობის საშინელი განცდა მქონდა და წარმოვიდგენ როგორ შეიძლება იტანჯებოდნენ ის ადამიანები, ვისაც ახლა შეხვდათ ვირუსი, როცა გარეთ 40 გრადუსია... სრულებით უნერგით და ქმედუნარო ვიყავი, მთელი კოვიდის პერიოდი გავატარე ძილში. მქონდა ჩუმი შიშები, რომ უცბად თრომბი განვითარდებოდა”*
(“...April is not as hot a month as July, but I had a terrible feeling of lack of air and I can imagine how people who have just met the virus when it is 40 degrees outside can suffer ... I was completely inactive and incapacitated, I spent the whole period of the infection asleep. I had silent fears that a blood clot would develop suddenly.”)
 (<https://www.radiotavisupleba.ge/a/31375013.html>)

The results of the present research reveal that one of the most frequent linguistic strategies to communicate threat to the readers of the Georgian online news is conditional sentences that, in terms of frequency, follow personalized stories (Example 12):

- (12) *თუკი კოვიდი უკვე გადაიტანეთ, არ იფიქროთ, რომ სამუდამო იმუნიტეტი გაქვთ და აცრა აღარ გჭირდებათ*
(If you have already had Covid, do not assume that you have permanent immunity and no longer need to be vaccinated.)
 (<https://www.radiotavisupleba.ge/a/31375013.html>)

The Georgian online news on the COVID-19 often apply statistics to communicate threat about the infection to the recipients. In this regard, it has to be mentioned that the key message the statistics presented in the Georgian online news reports convey deals with clear intention to provide evidence on positive outcomes of vaccination. Moreover, in the Georgian online news reports, statistics and personalized stories are combined and thus create a twofold effect on the recipient. In other words, it can be claimed that these sections of the Georgian texts represent more statistics within a personalized story than the standard type of statistics. What is more, the form and content of the statistics in the Georgian online news reports is similar to those applied in the English online news reports; news reports in both languages communicate positive content regarding vaccination, while negative figures are used to prove vulnerability of unvaccinated people (Example 13):

- (13) *გარდაცვლილთა 1%-იც არ იქნება აცრილი! ჩვენთან ერთი გარდაიცვალა მხოლოდ და ისიც ძალიან მძიმე თანმხლები დაავადებით - შეიძლება კოვიდის გარეშეც რომ გარდაცვლილიყო!*
(Even 1% of the dead might not have been those vaccinated! There was only one death case in our department and he/she also had severe accompanying diseases – he/she could have died without Covid too!)
 (<https://www.radiotavisupleba.ge/a/31375013.html>)

The figures obtained through the analysis of the tagged texts of the Georgian online news reports demonstrate that ontological metaphors are also used to some degree to communicate threat of the virus; however, regarding the frequency of application this stylistic device is less applied in Georgian texts than in English ones (Example 14):

- (14) “...ფსიქოლოგიურადაც განგრევს ეს მდგომარეობა, როცა არ ხარ მიჩვეული ... და ავადმყოფობის შემდეგ ძველ ქოსურ გარემოში გამოდიხარ, უცებ ხვდები რომ ორგანიზმი ვერ მოგყვება, თავბრულც გეხვევა და ვერ ფუნქციონირებ როგორც აქამდე.”
 (“...this condition also destroys psychologically when you are not used to it... and **after the illness you are in an old chaotic environment, you suddenly realize that your body can not follow you, you feel dizzy and cannot function as before.**”) (https://www.radiotavisupleba.ge/a/31375013.html)

Finally, in comparison to the English online news, the Georgian online news hardly ever use epithets to communicate threat to the recipient.

5.3 Comparative analysis of linguistic means in English and Georgian online news reports

The data presented in Table 1 below enable to obtain a comparative picture of the tendencies in utilization of threat markers in English and Georgian online news reports. The numbers provided separately for the power of each identified linguistic techniques provide quite a transparent ground for drawing conclusions separately for each language as well as detecting similarities and differences between them.

Linguistic markers	Metaphor	Authorities	Comparison	Statistics	Conditional	Cases	Epithet	(DAT)
AP English	0.8	0.6	0.7	0.1	0.02	0.1	0.1	0.1
AP Georgian	0.4	0.3	0.3	0.1	0.1	0.1	0	0.1

Table 1: Average power of linguistic markers in the Georgian and English news reports

The first horizontal row shows the types of linguistic techniques used in the English and Georgian online news to communicate threat and persuasion to the target audience, while the other two rows show the AP of linguistic techniques in the English and Georgian news

reports. The present table also allows observing the density (DAT) of each linguistic technique to communicate threat.

The below presented diagram created with the help of a corpus-based approach provides a valuable picture of specifics in application of various strategies to communicate threat and persuasion in English and Georgian online news on the COVID-19.

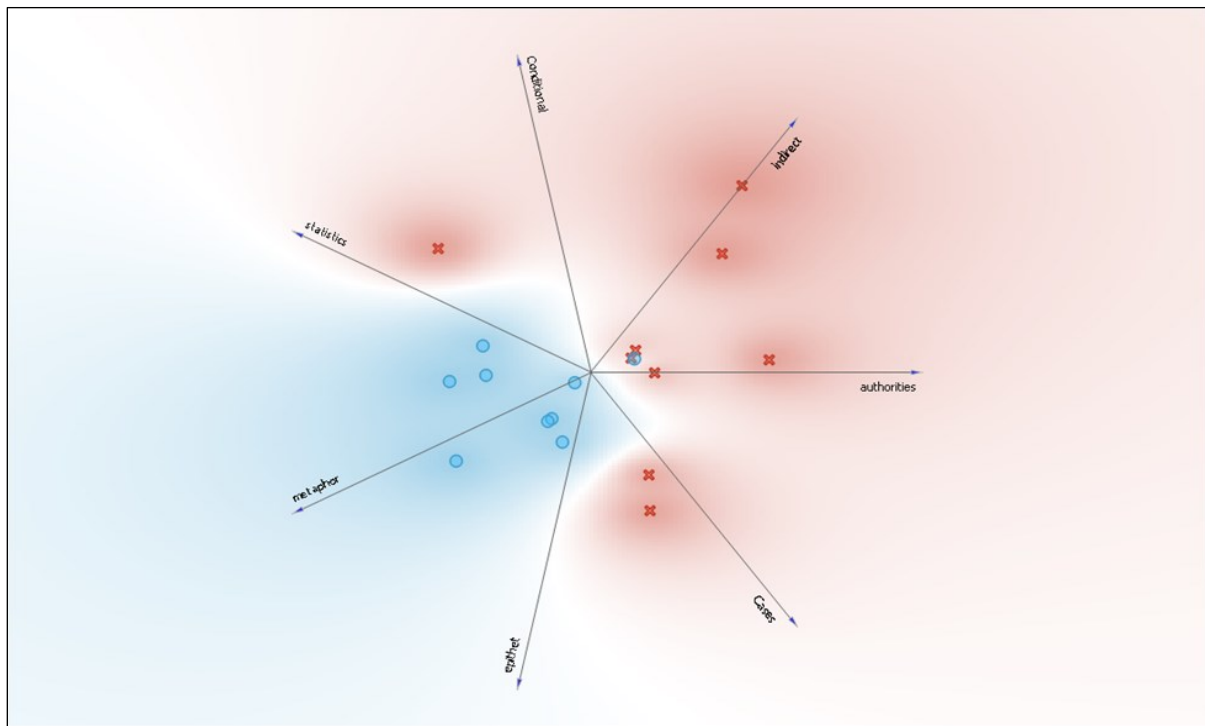


Figure 1: Radial diagram

The upper side of the radial diagram with the symbols of crosses shows the nature of utilization of various rhetorical strategies to communicate threat and persuasion in the English online news reports. Dispersion of crosses suggests that English online news reports apply a variety of linguistic techniques to communicate threat to their recipients. This means that the texts in English online news reports consider different ways of perception of information by the recipients. The more types of rhetorical strategies are used, the more possibility there is to reach the goal in terms of the desirable influence on the target audience.

The symbols of circles in the radial diagram represent the pattern of application of rhetorical strategies to communicate threat and persuasion in the Georgian online news reports. The pattern of location of the circles, which clearly shows much less dispersion, can be interpreted by the similar nature of linguistic techniques applied in the Georgian online news reports. Closeness of circles within the space for the Georgian online news reports

analysis shows that the Georgian online news reports may have been worked out by one group of people teamed up under the same ideology or vision. The fewer types of powerful tools to communicate a message are used, the less potential there is to get a favorable response from the recipients, and this can be proved by the number of the vaccinated people in Georgia, which is much lower in comparison with many other countries where more trust between the source and the recipient of the information was built.

6 Conclusion

The results of the research show the most frequently utilized linguistic techniques and strategies to convey the message on threat in the Georgian online news reports. Pursuant to the data given in the table demonstrating the numbers for each identified linguistic techniques in the Georgian online news reports, the most dominant linguistic technique targeted at communicating threat to the recipient is citation of authorized people or organizations. What is more, these names include local as well as internationally recognized people or institutions. The results of the research prove that comparative statistics for vaccinated and unvaccinated people is the second most widely applied persuasive technique used in the Georgian online news reports. The analyzed texts also demonstrate that telling personal stories about victims as well as survivors has a great impact on the emotions of the recipients. Moreover, direct citation of those who have experienced the unmerciful nature of the virus is a very powerful tool to motivate the target audience. Conditional sentences are the next linguistic techniques after the techniques of telling stories according to frequency of their application in the Georgian online news reports. The next widely applied linguistic strategy to communicate threat about the Corona virus in the Georgian online news reports is statistics, which is followed by metaphors as the least frequently utilized linguistic technique to communicate threat and persuasion to the recipient of the Georgian online news reports. The figure (0.9) presented in the very last row of the table shows the density of threat-invoking linguistic techniques in all analyzed sources of the Georgian online news reports. This number shows spatial extension of linguistic techniques within all Georgian texts selected for the corpus analysis.

The analysis of the English online news on the COVID-19 with the objective to identify certain persuasion strategies in the texts manifests the application of a variety of persuasive techniques to communicate the message on the killing virus. The data obtained from processing the English online news have enabled to observe the frequency of application

of each threat-conveying linguistic technique. Based on these data it can be stated that the message on the virus is communicated in two ways – direct and indirect. Moreover, the numbers suggest that the indirect techniques largely outweigh direct techniques used to persuade a recipient out of which statistics is the most frequently applied technique followed by description of scenes showing privileges for vaccinated people. Based on the results of the research, metaphors are the third most widely applied persuasive strategies in the English online news reports, while the strategy of naming authorities or authorized institutions stands between metaphors and the strategy of telling stories. From the point of view of frequency of application, conditional sentences follow the strategy of telling stories about the victims. The last row of the table shows that epithet is the least frequent rhetorical strategy to communicate threat or persuasion.

As mentioned above, the corpus-based analysis of the English online news reports revealed density of all linguistic techniques within all texts processed in the present research. The figure (0.4) represents correlation of the volume of all threat expressing techniques which, based on the theories presented above, contribute to persuading the audience, and the amount of all online news analyzed.

In conclusion, we can assume that Georgian news reports prefer utilization of fewer types of techniques to communicate threat to recipients and persuade them in comparison to the English online news that demonstrate utilization of a variety of linguistic techniques for the same purpose.

References

- Aristotle (1939) *The “Art” of Rhetoric: With an English Translation by John Henry Freese* [translated by J. H. Freese]. Harvard University Press.
- Beck, K. H. and Frankel, A. (1981) ‘A conceptualization of threat communications and protective health behavior.’ *Social Psychology Quarterly* 44, 204–217. <https://doi.org/10.2307/3033834> [PubMed] [CrossRef] [Google Scholar]
- Buzan, B., Waeber, O. and de Wilde, J. (1998) *Security: A new Framework for Analysis*. Boulder, CO and London, UK: Lynne Rienner Publishers. [GoogleBooks]
- Festinger, L. (1957) *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press. <https://doi.org/10.1515/9781503620766>
- Fisher, W. R. (1987) *Human Communication as Narration: Toward a Philosophy of Reason, Value and Action*. Columbia, SC: University of South Carolina Press.
- Hovland, C., Janis, I. and Kelley, H. (1953) *Communication and Persuasion*. New Haven, CT: Yale. [Google Scholar]
- Kozlova, T. (2021) ‘Cognitive metaphors of the COVID-19 pandemic in business news.’ Published online by EDP Sciences. *IV International Scientific Congress proceedings “Society of Ambient Intelligence – 2021” (ISCSAI 2021)*. <https://doi.org/10.1051/shsconf/202110002004>
- Lazarus, R. S. (1991) *Emotion and Adaptation*. New York, NY: Oxford University Press. [Google Scholar]

- Mongeau, P. A. (2013) 'Fear appeals.' In: Dillard, J. P. and Shen, L. (eds) *The Handbook of Persuasion*, 2nd ed. Thousand Oaks, CA: Sage. 184–199. <https://doi.org/10.4135/9781452218410> [CrossRef] [Google Scholar]
- Neville Li, Ch. H. (2020) 'How to construct Covid-19 as a national security threat in public discourse?' Lingnan University. Viral Discourse Health Communication Network. [ResearchGate]
- O'Keefe, D. J. (1990) 'The study of persuasive effects.' In: O'Keefe, D. J. (ed.) *Persuasion: Theory and Research*. Newbury Park, CA: Sage. 169–180 [Google Books]
- Perelman, C. and Olbrechts-Tyteca, L. (1969) *The New Rhetoric: A Treatise on Argumentation*. Notre Dame: University of Notre Dame Press. [WorldCat]
- Petty, R. E. and Cacioppo, T. J. (1986) *The Elaboration Likelihood Model of Persuasion. Advances in Experimental Social Psychology*. [Researchgate] [https://doi.org/10.1016/S0065-2601\(08\)60214-2](https://doi.org/10.1016/S0065-2601(08)60214-2)
- Rogers, R. W. (1983) 'Cognitive and physiological processes in fear appeals and attitude change: A revised theory of protection motivation.' In: Cacioppo, J. T. and Petty, R. E. (eds) *Social Psychophysiology: A Source Book*. New York, NY: The Guilford Press. 153–176. [Google Scholar]
- Simons, H. W. (1976) *Persuasion: Understanding, Practice, and Analysis*. Reading, MA: Addison-Wesley Pub. Co.
- Witte, K. (1992) 'Putting the fear back into fear appeals: The extended parallel process model.' *Communication Monographs* 59(4), 329–349. <https://doi.org/10.1080/03637759209376276> [CrossRef] [Google Scholar]

Sources

- Burduli, A. (2021, July 24). კოვიდი ორჯერ ზედიზედ-რატომ არის მნიშვნელოვანი აცრა ინფიცირების შემდეგ (*Covid twice in a row - why is it important to get vaccinated after infection*). საზოგადოება (community). <https://www.radiotavisupleba.ge/a/31375013.html>
- წვრილშვილიანი ოჯახი დარჩა უმამოდ...აცრილი რომ ყოფილიყო აუცილებლად ცოცხალი იქნებოდა-რას წერს ბოჭორიშვილის კლინიკის დირექტორი? (*A family with a small child was left without a father... If he had been vaccinated, he would definitely be alive - what does the director of the Bochorishvili clinic write?*). (2021, December 8). Ambebi. Ge. <https://www.ambebi.ge/article/269360-cvriilshviliani-ojaxi-darcha-umamod-acrili-rom-qo/>
- Gallaghet, J. (2021, October 22). *Covid: Why is coronavirus such a threat*. BBC News. <https://www.bbc.com/news/health-54648684>
- Holcombe, M. (2021, September 1). *CDC asks the unvaccinated not to travel this weekend and says even vaccinated need to weigh the risk*. CNN. <https://edition.cnn.com/2021/09/01/health/us-coronavirus-wednesday/index.html>
- John, T. (2021, September 1). *100,000 more could die from Covid-19 in US by December*. CNN. Coronavirus: fact VS. Fiction. <https://edition.cnn.com/2021/09/01/world/coronavirus-newsletter-intl-1-09-21-duplicate-2/index.html>
- Marquez, M. and Yan, H. (2021, December 10). *Covid -19 patients at this hospital are dying 'at a rate we've never seen die before'—and it's taking a toll on health care workers*. CNN. <https://edition.cnn.com/2021/12/09/us/hospital-covid-19-deaths-michigan/index.html>

Zeinab Gvarishvili is Associate Professor of English Linguistics at Batumi Shota Rustaveli State University, Batumi, Georgia. She specializes in diachronic studies of English language, pragmatics, discourse analysis and comparative linguistics.

Address: Zeinab Gvarishvili, Department of European Studies, Specialty of English Philology, Batumi Shota Rustaveli State University, 35 Ninoshvili St., Batumi 6010, Georgia. [e-mail: zeinab.gvarishvili@bsu.edu.ge]

Nana Mazmishvili is Associate Professor of English Linguistics at Batumi Shota Rustaveli State University, Batumi, Georgia. She specializes in translation studies, pragmatics/linguistics, discourse analysis and applied Linguistics.

Address: Nana Mazmishvili, Department of European Studies, Specialty of English Philology, Batumi Shota Rustaveli State University, 35 Ninoshvili St., Batumi 6010, Georgia. [e-mail: nana.mazmishvili@bsu.edu.ge]

ADVERSE ‘NATIVE SPEAKER’ EFFECTS ON ANGLOPHONES IN THE MULTILINGUAL WORKPLACE

Veronika Lovrits

<https://doi.org/10.5817/CZ.MUNI.P280-0212-2022-4>

Abstract

In order to explore the social effects of the differentiation between ‘native’ and ‘non-native’ varieties of English, the present qualitative study followed six Anglophone trainees in an EU institution in Luxembourg. Data were gathered in 2018/19 and 2020/21, combining on-site observations with longitudinal and one-off interviews. Research incited participants’ reflections on language practices at work and mapped their discursive positioning. Conclusions drawn from a sociolinguistic analysis of stances show that, despite the vagueness and lack of clearly definable linguistic characteristics for the category, being labelled as ‘native English’ brought distinct negative effects to the experience of the ‘native English’ participants. The contribution highlights the social constructivist character of the native/non-native dichotomy and draws attention to its adverse effects in the multilingual workplace and beyond.

Keywords

multilingual, sociolinguistic, native, English, language ideology, EU, stance

1 Introduction

The research presented in this contribution took place in the linguistically superdiverse context of Luxembourg. Although Luxembourg counts only about 630,000 inhabitants, its economy attracts strong labour force migration and the quotidian is traditionally multilingual in Luxembourg (Purschke 2020). Nearly 200,000 cross-border workers commute to Luxembourg from the neighbouring countries (STATEC 2020) and the state administration can be addressed in four languages: Luxembourgish, French, German, and German sign language. In addition, English has gained more and more ground in the public space (Horner & Weber 2011). While the reasons for the rise of English use in Luxembourg’s multilingual public spaces have not been the focus of existing research, part of the explanation seems to be grounded in language ideologies accompanying the use of English as *lingua franca* or language of international business (de Bres 2017).

Nonetheless, the aim of the presented research is neither to exoticise Luxembourg as an *élite* multilingual space in the European Union, nor to explore the reasons for the continuing spread of English in global urban spaces with high migration. The objective is to investigate the social effects of the social distinction between ‘native’ and ‘non-native’ varieties of English in a linguistically diverse environment. The main question will be posed from a critical sociolinguistic angle and the contribution will draw attention to the problematic character of the ‘native/non-native’ dichotomy: “What social effects does the ‘native English speaker’ position bring about in a multilingual team of a supranational institution?”

The coming section of this paper will briefly introduce the place of research, participants and data collection. The next section will conceptualise the ‘native speaker’ and describe the method of sociolinguistic stance analysis. Then, the results will be presented: the typical stance positioning of the ‘native English speaker’ and the communicative hiccup in which the ‘native English’ positioning led to problematic language behaviour. This will be followed by a discussion on the relevance of the findings and implication for practice. Finally, a conclusion section will complete the contribution.

2 Place of research and data collection

The present analysis is part of a study investigating reflections on language practices in a terminology and communications team in an EU institution in Luxembourg. The unit did not represent one EU member state or language. Instead, it served the institution through whole and coordinated inter-unit-, inter-institutional-, and external cooperation projects. The unit was officially multilingual. However, its members used English as the main vehicular language to address a transnational (primarily namely European) public.

Data collection followed the principles of participative action research (McIntyre 2008) and employed a mix of qualitative methods – on-site observations, reflective drawing, longitudinal-, and one-off interviews. Interviewing incited participants’ reflections on language practices and socio-linguistic discursive positioning of the team members, proceeding in iterative rounds. After each round, the researcher shared her pre-analyses with the participants, who were then free to discuss the topics before the next round of interviews.

This contribution focuses on six Anglophone trainees during their five-month long traineeship. All of them were discursively positioned as ‘native English speakers’ in the team in the pilot phase in 2018/19 and the follow-up phase in 2020/21. Both phases started with the trainees reflecting on a traineeship term that was about to pass or had already passed

(Ben – Irish, Lucy – British, Kate – Maltese and Jolene – Luxembourgish) and continued by rounds of in-depth interviewing during their current traineeship term (Florence – British and Valerie – German). The participants’ names were pseudonymised to retain the participants’ anonymity.

The pilot study focused on the Anglophone experience and a report discussing the effects of language ideologies from a macro-social perspective has since been published in the *Journal of Sociolinguistics* (Lovrits & de Bres 2020). A follow-up study then involved all members of the team and explored the broader topic of languages and language practices in the unit. The position of the ‘native English’ trainee was not central to the follow-up investigation, but the ‘native English’ topic still resonated in the team and marked a pattern in the workplace. Thus, the present contribution will recall the main findings of the pilot project in light of additional data from the follow-up project, while further discussing the relevant theoretical aspects that may enhance methodology and practice in linguistics research.

3 Social construction of the ‘native’

The main interpretative framework will be set using the social constructivist paradigm that focuses on the socially constructed aspects of the commonly perceived reality (Berger & Luckmann 2001). Bourdieu’s poststructuralist social theory (1998, 2001) will further help to explain the social effects of the ascription of ‘native’ or ‘non-native’ status. This will, in turn, clarify that the potential linguistic difference between the ‘non/native’ varieties may be of less significance than the social ascription of the symbolic sociolinguistic status to its speakers.

According to Bourdieu (1998), individuals in the society take positions in an abstract matrix of mutual relations based on two basic principles – distinction and differentiation. Sometimes, following the interests of powerful social groups, the socially constructed differentiations are treated as innate, given, unchangeable distinctions. This limits the discussion on the fairness or practicality of the social set up, as it is considered a given. According to Bourdieu (2001), language hegemony produces a codified language and links it to a political entity. The standard language becomes officially produced, unified, checked, tested, and verified by the state and its certified institutions. Language hegemony then allows for only one ‘right’ variety; the geographical and social alternatives are demoted for the purpose of controlling their users (ibid.: 70-71).

The idea of a named language without varieties is mirrored in what is theorised as stereotypical *beliefs* (Doerr 2009: 20) and *language ideologies* (Pennycook 1994: 176) based

on the notion of ‘native speaker’. Within those beliefs/language ideologies, the ‘native speakers’ are seen as having a linguistic competence inherent to a homogeneous population in the given national state. This is portrayed through references to the national culture and language, disregarding the differences in socioeconomic stratification which are mirrored in the varieties of language use (Rampton 2003). The ‘native’ superiority is a perception informed by a monolingual lens that overemphasises the separation of languages (Arocena, Cenoz & Gorter 2015) and sees named languages as a mythical homogeneous entity without variation. However, no variation is intrinsically better than another; it is the status of their users that makes one variation more desirable than another (McKenzie 2013). Moreover, the ‘native speaker’ notion has troublesome links to colonialism (Pennycook 1994). Piller (2018), namely, shows how it has been used to discursively denigrate or elevate certain social groups in the (colonial) past.

The allegedly inevitable character of the ‘native’ category is also discursively strengthened by an imagined biological link. The term ‘mother tongue’ is often used as a synonym for the ‘native language’, accentuating the allegedly inborn nature of the ‘native’ category (Davies 2006). Drawing on the German philosophical tradition, Yildiz (2012: 12-13) describes this link as a “historic family romance with the mother tongue”, a fantasy, that the mother tongue is biologically connected to the maternal body.

Linguistic research and theory have also demonstrated that the ‘native speaker’ category is too vague a notion for research, being ultimately decided by a (self-)ascription of the ‘non-/native’ membership with a reference to (auto-)biography (Davies 2006). Recent research has also pointed out that “methodological nationalism” (Kraus 2018: 93) limits researchers’ understanding of language use in real life by associating language with (national) state. The critical approach has been gaining acceptance in the global research community over the last few decades. For instance, the editorial guideline of the *Journal of Multilingual and Multicultural Development* now encourages authors “to reject a deficit view of multilingualism and hence avoid terms such as ‘native speaker, non-native speaker’ and use neutral terms such as ‘L1 user, LX user’ instead”.

Despite the above trends, the notion of ‘native’ speaker or language often remains unscrutinised in language research. Explorations of the social effects of its ideological connotations have been mainly limited to the domain of education, where the ‘non-native’ teachers face systemic discrimination in the global labour market (Doerr 2009) and learners have to bow down to the pedagogically questionable ideal of the ‘native speaker’ (Dewaele &

Saito 2022). Thus, this contribution elaborates on the constructivist character of the topic and the respective agency that individual speakers have over its effects, pointing out the importance of critically examining the use of the ‘native’ category in research and practice. The expressions ‘native’ and ‘non-native’ will be used in quotation marks in order to remind the reader that the category is treated as socially constructed and imagined by participants in varying ways.

4 Sociolinguistic analysis of stance

Stance as an analytical tool is defined according to three constitutive aspects. It consists of the object of stance, stance positioning (of the speaker towards the object), and its alignment or misalignment with the positioning of other speakers (Du Bois 2007). Stance is considered to be the speaker’s primary concern in conversation (Kiesling 2009) and has been employed in linguistics in two main ways.

First, interactional linguistics uses stance to focus on the use of linguistic forms, words or syntactic forms indicating conversational turns (Kärkkäinen 2003, Wu 2003, Du Bois 2007). Second, discourse analysis links macro-social (often political) discourse to the micro-level of individual and group interests through stance (Englebretson 2001, Jaffe 2009). So, the discursive take on stance analytically joins the ideologically (discursively) shaped *structure* of workplace communication with the individual *agency* of speakers over their communicative status (Jaffe 2009).

The presented study aims to link individual experience to the phenomena of broader social importance, notably language ideologies and their socio-cultural context. Thus, the discursive analysis of stances will be employed. The next subsection will present the results of the analysis, starting with the identification of the common object of stances – who is considered the ‘native English speaker’, his/her typical stance positioning, and the social effects it brings about. In the second subsection, an analysis of a recurring tension linked to language choice will show how the positioning of ‘native English speaker’ created tensions.

4.1 Typical positioning of the ‘native English speaker’

The investigated object of stance was the ‘native English speaker’. Participants treated the ‘native speaker’ category as a fact defined by the geopolitical affiliation of the speaker. On top of that, the category was culturally defined as ‘British’ for the British participants, while the mental representations of other participants included other potential English-speaking environments in Europe. The bilingual participants also felt less constrained by the

fact that they were considered ‘native English speakers’, because their positioning was less linked to their identity, as Kate’s stance exemplifies:

- (1) *K: So- (.) bu- but ahm, being referred to kind of, as a- as a native English speaker was- was kind of a (..) nice thing for me. I- I would put it. Yeah, I- it was a nice thing, because it was- first of all, it wasn’t something I expected or applied to do.*

Along with the geopolitical affiliation, the ‘native English speaker’ was identified by the ability to not make “huge/terrible/detrimental linguistic mistakes”. Nonetheless, over the course of five traineeship terms, no participant could recall any such a mistake. At the same time, the ‘native English’ were praising their colleagues’ English, as illustrated below¹:

- (2) *V: Sometimes they’re not too sure about their English and it’s really – it’s top-notch. It’s really good.*

Researcher: But you keep proofreading anyway.

V: Yes, because of course it has to be the .. top standard.

Although the participants did talk about coming across texts that “apparently were not written by a ‘native’”, non-idiomatic texts never caused any problem. It is remarkable that the study could not trace any experience of a problem with a ‘non-native’ variety of English, even though there were, at times, no ‘native English’ trainees on the team. Instead, the participants focused on imagined social consequences, such as being laughable or making the unit look bad. The ‘native English’ trainee was positioned as a protector against imagined reproaches concerning non-idiomatic English use. His/her main role was proofreading. That positioning was both advantageous and limiting.

First, some ‘native English’ participants enjoyed proofreading, whereas others frowned upon it. In any case, they had not expected to be given this task when they applied for the traineeship. Moreover, proofreading could take up a significant share of their time at work, which was not controlled by the superiors. Indeed, the actual scope of ‘naturally occurring’ proofreading tasks was one of the surprises this study brought to the light.

Second, the ‘native English’ trainees coming from English-speaking countries understood that the ‘native speaker’ category was defined by extralinguistic aspects (being from somewhere or born to somebody who is from somewhere), not by their actual linguistic performance. As such, they did not consider their ‘native language’ to be a personal skill they could showcase at work. As a result, they felt objectified. They were under the impression that the ‘native’ language variety was considered important, as opposed to their individual skills. Ben gave a sigh that summed up this type of positioning:

- (3) *B: ... Uhm, I didn't really know it beforehand (..) but it kind of became clear on the first day (..) that (.) I was selected because of the fact that I was a native speaker. So (..) I don't really know how I felt about that at the time? Because I'd like – I'd like to think that it was (..) my experience and my studies, my (..) professional experience to that date.*

The third limiting consequence of the 'native English speaker' positioning was that the 'native English speakers' were considered monolingual and therefore found themselves in an undesired English-only communicative environment. Indeed, the 'native English' variety was put at the top of the hierarchy of English uses, but the trainees' ultimate goal was to become actively *multilingual*. Thus, the English-only regime expected of the 'native English speakers' thwarted their desire to develop multilingual skills while working abroad. They would have preferred more opportunities to practise other languages:

- (4) *F: I proofread English and (.) it's needed (.) and yeah, so I think that it does the work, yeah. I think I am very much in my comfort zone, yeah. As well. It's not always something I (..) really enjoy being. I think it's good to be, you know, not always in your comfort zone, but – (..) but yeah.*

The Luxembourgish trainee was an exception. She was considered 'native English' based on one of her home languages and the fact that she had taken her university studies in the UK (she was also not 'abroad' in Luxembourg). Her social life was already multilingual, so she did not mind the English-only workplace (she also chose to use Luxembourgish in her interview, below is the researcher's own translation):

- (5a/Luxembourgish) *J: Et si Leit, di mengen, ech si vläicht- ech wär e Native-Speaker? Op Englesch. Well wann ech Englesch schwätzen, hunn ech en- hunn ech engleschen Accent? An ech mengen, datt erkläert firwat d'Leit mech heiansdo froen, wéi laang ech schonn a Lëtzebuerg gewunnt hunn? (lächelt) Um (...) Es sinn den déi Alles muss iwwerliesen? An ech maaen dat och gär.*

- (5b/English) *J: There are people who think I am – I am a native speaker? In English. Because when I speak English, I have an English accent? And I think that explains why people sometimes ask me, how long I have been living in Luxembourg? (smiles) Um (...) so I am the one who has to proofread everything? And I like to do that, too.*

The above excerpt also draws attention to the tendency to assume that a 'native' speaker must have a stable geopolitical affiliation and biography, as opposed to a migration or multilingual background. Such methodological nationalism obscures the criteria that define what language performance from a given geo-political area can be considered the 'native' variation for a linguistic corpus. Nevertheless, this phenomenon has hardly received any critical attention.

4.2 ‘Native English speaker’ in the middle of tensions

In general, the trainees enjoyed their stay and built positive relationships within the team. However, the analysis showed that the discursive elevation of the ‘native’ performance by superiors incited a tense atmosphere amongst junior colleagues, fuelling competition between the linguistic repertoires and language ideologies in the society to the detriment of team cooperation. The ‘native English speaker’ positioning brought about unchallenged presumptions of power relations that twisted their communication.

The British participants strongly felt their ‘native English’ positioning – they both referred to it when reflecting on how they dealt with the language choices of their colleagues. At some point in the pilot project, dialogues in the trainee office turned to Spanish, which the British trainees did not understand. Data and theory suggest the following explanations for this change (a combination of these may be relevant in this case). First, the follow-up study revealed that some team members did not realise how uncomfortable it may be, to be present in a room while one’s co-workers are speaking in an unfamiliar language. One reason for the Spanish-speaking colleagues using Spanish could be that they were immersed in the discourse on ‘global Spanish’ as a language that is somehow, at least passively, understood ‘everywhere’ (van den Worp et al. 2018). Over time, the use of Spanish could have developed into a more or less conscious act of revenge on the ‘native English speakers’.

Whatever the reason for the use of Spanish, both British trainees felt excluded, by way of a direct discursive reference to their ‘native English’ positioning in the workplace. However, they were reluctant to ask colleagues to speak English, assuming that this would reinforce the stereotype of privileged monolingual British people abroad. Lucy tried to seek support from another Spanish colleague, complaining about her exclusion. Instead, she received confirmation that her stereotypical positioning as an English-speaking monolingual was salient:

- (6) *L: I couldn’t understand, like I couldn’t work on the project, because I didn’t know what they were talking about. And I was complaining to [a colleague] and he was like: “Oh, you English speakers!” I understand – (..) But I was like – it’s – you know, if it’s for work and I couldn’t understand what we’re working on?*

While Lucy gave up and the colleagues finished the project without her, Florence (in another team of trainees) was more active when she was similarly excluded. When her colleagues started to speak Spanish in her presence, she hinted indirectly, and then directly, that the situation was uncomfortable for her. Since this seemed to have no effect, she turned to passive-aggressive behaviour and “tactically” forced the colleagues to speak English, in her

own words. She arranged a loud talk to her colleagues and made sure their superiors could hear it:

- (7) *F: [I will-] (bursting out laughing) Like it wouldn't be outwardly, but I'll be keeping it in mind. Like: "Okay!" (amused voice) (..) You know. I'll play up to (..) my native English role (amused voice) you know (..) if I need to. (laughing)*

That strategy technically worked, since the dialogues turned to English again. However, she still felt excluded by her colleagues.

Although the collected data do not allow for a conclusion that the 'native English' were excluded by their colleagues on purpose, trainees in the follow-up project in 2020-21 did cautiously talk about "leveraging" the 'native' privilege with skills in other languages, which the 'native English' trainees supposedly lacked. That said, the bilinguals from the follow-up project did not share the experience of 'native English' exclusion. However, their 'non-native English' colleagues perceived the adulation of 'native English' as an injustice and wanted the seniors to either refrain from elevating the 'native English' or to recognise other relevant skills in the 'non-native English' trainees with a similar discursive gratitude.

5 Discussion

Bourdieu notes (1998) that declaring categories and assigning them to a group of people engenders social relations which are often uncontested, because they are gradually perceived as 'natural' and a given. This is the case of the 'native speaker' and 'native language' notion. Had the 'native English speakers' in this study discussed the social constructedness of sociolinguistic positioning during their studies, they would no doubt have been better prepared for a multilingual workplace. They could have, at least, tried to question the stereotypes or expressed their desire to be included and/or to practise other languages.

To overcome the illusion of inevitability of the 'native' categories, we may need to critically rethink their use, not only in our day-to-day lives but also in professional discussions. In cases where the localisation and socio-historical context of language is relevant, we should clearly define what constitutes the linguistic variation we are looking for. Instead of referring to the vague 'native speaker' category, we may want to ask: what is an acceptable deviation in what aspects of language use (vocabulary, grammar, style, intertextuality, etc.) and what are the (social, political, legal, professional and other) aspects we have to respect?

In applied linguistics, meta-pragmatic awareness and framing of varieties in the classroom should also be put under the spotlight. Although language teachers must choose

a standard model for teaching, they could prepare students for more than the obligatory standardised tests, for instance by considering life in multilingual environments. Indeed, practitioners sometimes get the impression that teaching language in laboratory-like conditions is enough, because all that the public can see are the results of standardised tests (CEFR/TOEFL). However, this teaching method disregards the social function of language, the individual and social identity linked to the sociolinguistic positioning, and the social and historical (in the case of English post-colonial) contexts in which language users must navigate worldwide.

Ultimately, in linguistically diverse contexts of international mobility, it may be difficult to assume language competence according to country, language of schooling or ‘home languages’. Even the ‘native’ category, defined as a monolingual experience from a defined geo-political area, does not guarantee that the person uses a locally typical or standard language, let alone that s/he will be ready for professional linguistic performance. The language use of the ‘natives’ varies as much as that of the ‘non-natives’, mirroring their socioeconomic, professional or family background, situational aspects and personal ability to use language in a pragmatically effective way.

6 Conclusion

Two main issues have been highlighted in this contribution. The first is the unnecessary ‘native English speakerism’ accompanying language production intended for an international public. Despite language performance being at the centre of all activities in the terminology and communications unit, the potential linguistic differences between the varieties did not matter in this workplace as much as the social ascription of the symbolic sociolinguistic status. The ‘non-native’ varieties were discursively constructed as socially risky, while the ‘native English speakers’ were assigned a symbolic power, a ‘magical touch’ that was supposed to protect the unit from external criticism. In the time spent researching this team, no instances of criticism were uncovered – and it is likely that these reproaches never occurred, since all team members were language professionals working in a global context. Nevertheless, even though the risk of non-standard linguistic fails was apparently elusive, the social effects of the ‘native English speaker’ positioning were experienced as real.

This brings us to the conclusion of the second main issue. The ‘native English’ participants were ascribed a workplace role defined by their sociolinguistic membership which was traditionally framed as advantageous but was not always so in practice. Not only were the ‘native English’ speakers often disappointed by the social effects of their

sociolinguistic positioning in the team, the study also showed that the discursive elevation of the ‘native’ English variety and ‘its’ speakers was at times to the detriment of the team’s ability to cooperate. The exact trigger(s) for such a situation are not fully clear. The tensions most probably arose when the cultural and language ideology aspects of the ‘native English’ positioning came to the foreground and/or when a subliminal competition for linguistic control over the workplace was triggered among junior team members.

As for the limitations, the presented findings stem from a micro-level qualitative study and in that sense, they are highly dependent on its context. However, this contribution illustrates a novel perspective that questions assumptions guiding research design in research on language. With the aim of enhancing research methodology, future research may benefit from paying more attention to the social effects of meta-pragmatic talk and engaging in critical investigations of conceptual boundaries of traditional heuristic tools like the ‘native speaker’.

To sum up, the identified adverse social effects of the ‘native English speaker’ positioning, an undesired English-only communicative space together with the sociolinguistic positioning burdened the ‘native English speakers’ with unwanted (often amateur) proofreading and thwarted their plans to develop multilingual skills while working abroad. Moreover, the discursive elevation of the ‘native speakers’ at times incited a tense atmosphere among the trainees. The conclusions align with the findings of critical research employing other methods over the last few decades – despite the vagueness, imaginary character, and no clearly definable *linguistic* characteristics constructing the ‘native’ category, its use can induce real-life ‘native speaker’ effects. However, it should be underlined that individual language users are not mere puppets in a scene set by social conditions. Social discourse creates a matrix of interpersonal power relations, mutual sociolinguistics positioning and prestige of certain varieties, but individuals always retain their agency.

Knowledge of the above offers an empowering socio-pragmatic awareness. Just as the awareness of malleability of sociolinguistic positioning and the related social effects seems indispensable for teachers and other language professionals, so is it also for students and the general public. It is moreover important for the interdisciplinary dialogue. We should acknowledge differing attitudes towards the variations, be sensitive to stances towards them, and in doing so, we may cultivate them in professional as well as informal discussions.

Notes

¹ Transcription conventions: extracts are transcribed broadly, with paralinguistic features and seconds of silence indicated in brackets, emphasised in bold. Interrupted sentence is indicated with a dash.

References

- Arocena, E., Cenoz, J. and Gorter, D. (2015) 'Teachers' beliefs in multilingual education in the Basque country and in Friesland.' *Journal of Immersion and Content-Based Language Education* 3(2), 169–193. <https://doi.org/10.1075/jicb.3.2.01aro>
- Berger, P. L. and Luckmann, T. (2001) *Sociální konstrukce reality. Pojednání o sociologii vědění.* (From English orig. *The Social Construction of Reality*). Praha: Centrum pro studium demokracie a kultury.
- Bourdieu, P. (1998) *Teorie jednání.* (From the French orig. *Raisons pratiques*, 1994). Praha: Karolinum.
- Bourdieu, P. (2001) *Le langage et le pouvoir symbolique.* (From the English orig. *Language and Symbolic Power*, 1991). Cambridge: Polity Press.
- Davies, A. (2006) 'The native speaker in applied linguistics.' In: Davies, A. and Elder, C. (eds) *The Handbook of Applied Linguistics*. New York, NY: Blackwell Publishing. 431–450. <https://doi.org/10.1111/b.9781405138093.2005.00023.x>
- de Bres, J. (2017) 'Friend or foe? The discourse of the rise of English in Luxembourg.' *Forum für Politik, Gesellschaft und Kultur in Luxemburg*. December 2017, 54–56.
- Dewaele, J. M. and Saito, K. (2022) 'Positive psychology can help overcome the pernicious native speaker ideology.' *The European Educational Researcher*. <https://doi.org/10.31757/euer.526>
- Doerr, N. M. (ed.) (2009) *The Native Speaker Concept: Ethnographic Investigations of Native Speaker Effects*. Berlin, Germany: Walter de Gruyter. <https://doi.org/10.1515/9783110220957>
- Du Bois, J. W. (2007) 'The stance triangle.' In: Englebretson, R. (ed.) *Stancetaking in Discourse. Subjectivity, Evaluation, Interaction*. Amsterdam, Netherlands: John Benjamins. 139–182. <https://doi.org/10.1075/pbns.164.07du>
- Englebretson, R. (ed.) (2001) *Stancetaking in Discourse*. Amsterdam/Philadelphia: John Benjamins.
- Horner, K. and Weber, J. J. (2011) 'Not playing the game: Shifting patterns in the discourse of integration.' *Journal of Language and Politics* 10(2), 139–159. <https://doi.org/10.1075/jlp.10.2.01hor>
- Jaffe, A. (ed.) (2009) *Stance: Sociolinguistic Perspectives*. New York, NY: Oxford University Press.
- Kärkkäinen, E. (2003) *Epistemic Stance in English Conversation: A Description of its Interactional Functions, with a Focus on I think*. Pragmatics & Beyond, New Series 115. Amsterdam and Philadelphia: John Benjamins Publishing. <https://doi.org/10.1075/pbns.115>
- Kiesling, S. F. (2009) 'Style as stance: Can stance be the primary explanation for patterns of sociolinguistic variation?' In: Jaffe, A. (ed.) *Sociolinguistic Perspectives on Stance*. Oxford: Oxford University Press. 171–194. <https://doi.org/10.1093/acprof:oso/9780195331646.003.0008>
- Lovrits, V. and de Bres, J. (2021) 'Prestigious language, pigeonholed speakers: Stances towards the 'native English speaker' in a multilingual European institution.' *Journal of Sociolinguistics* 25(3), 398–417. <https://doi.org/10.1111/josl.12431>
- McIntyre, A. (2008) 'Participatory action research.' *Qualitative Research Methods* 52. Los Angeles: Sage. <https://doi.org/10.4135/9781483385679>
- McKenzie, R. M. (2013) *Changing Perceptions? A Variationist Sociolinguistic Perspective on Native Speaker Ideologies and Standard English in Japan*. Online document. https://www.researchgate.net/publication/261729625_Changing_perceptions_A_variationist_sociolinguistic_perspective_on_native_speaker_ideologies_and_standard_English_in_Japan/ref erences. Last accessed: 07. 06. 2022. <https://doi.org/10.21832/9781847698704-020>
- Pennycook, A. (1994) *The Cultural Politics of English as International Language*. London: Longman.
- Piller, I. (2018) 'How to end native speaker privilege. Language on the move.' 31 May 2018. Online document. 7 June 2022. <http://www.languageonthemove.com/how-to-end-native-speaker-privilege>

- Purschke, C. (2020) 'Attitudes toward multilingualism in Luxembourg. A comparative analysis of online news comments and crowdsourced questionnaire data.' *Frontiers in Artificial Intelligence* 3. <https://doi.org/10.3389/frai.2020.536086>
- Rampton, B. (2003) 'Hegemony, social class and stylization.' *Pragmatics* 13(1), 49–83. <https://doi.org/10.1075/prag.13.1.03ram>
- Spotti, M. and Blommaert, J. (2016) 'Bilingualism, multilingualism, globalization, and superdiversity.' In: García, O., Flores, N. and Spotti, M. (eds) *The Oxford Handbook of Language and Society*, Vol. 1. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190212896.013.1>
- STATEC. (2020) *Luxembourg in Figures 2020*. Luxembourg: STATEC.
- van der Worp, K., Cenoz, J. and Gorter, D. (2018) 'Multilingual professionals in internationally operating companies: Tensions in their linguistic repertoire?' *Multilingua* 37(4), 353–375. <https://doi.org/10.1515/multi-2017-0074>
- Wu, R. J. (2003) *Stance in Talk*. Amsterdam and Philadelphia: J. Benjamins.
- Yildiz, Y. (2012) *Beyond the Mother Tongue: The Postmonolingual Condition*. New York: Fordham University Press. <https://doi.org/10.2307/j.ctt13x0cqr>

Veronika Lovrits is a doctoral researcher in linguistics with previous professional experience in legal counselling and social inclusion mainstreaming. Her critical sociolinguistic research investigates language-related tensions in the multilingual workplace and pragmatic ways to address them.

Address: Veronika Lovrits, Research Institute of Multilingualism, Faculty of Humanities, Education and Social Sciences, Université du Luxembourg, Campus Belval, Maison des Sciences Humaines, 11, Porte des Sciences, L-4366 Esch-sur-Alzette. [e-mail: veronika.lovrits@uni.lu]

**HEDGING LIKE A PROFESSIONAL:
A CORPUS-DRIVEN APPROACH TO INTERACTIONAL
METADISOURSE IN ENGLISH LEARNER
ACADEMIC WRITING**

Markéta Malá

<https://doi.org/10.5817/CZ.MUNI.P280-0212-2022-5>

Abstract

The paper explores the phraseology of English academic texts written by Czech university students in comparison with English L1 novice and expert academic writing, focussing on hedging patterns. It combines contrastive analysis and learner corpus research, taking as its starting point recurrent multi-word patterns overused or underused by both groups of novice writers (e.g. *it can be said that* and *it seems* ADJECTIVE *that/to*, respectively), or by English L2 writers (e.g. *as it seems*) in comparison with L1 expert writers. The findings suggest that when expressing epistemic stance in their English academic papers, Czech university students have to face two types of challenge – the more prominent ‘academic’ challenge, i.e. the lack of academic experience, and, despite their proficiency, also the ‘linguistic’ EFL challenge.

Keywords

metadiscourse, hedging, novice academic writers, phraseology, learner corpora

1 Interactional metadiscourse, hedging and novice academic writers

It is hardly necessary to introduce the notion of metadiscourse today. In studies of academic writing, “the fact that, as we speak or write, we negotiate with others, making decisions about the kind of effects we are having on our listeners or readers” (Hyland 2005: 3) has received considerable attention with respect to variation across disciplines, genres, languages, language backgrounds, as well as language proficiency and academic status (e.g. Abdi 2002, Hyland 2004, Gray & Biber 2012, Bruce 2016, Lancaster 2016, Aull 2019, Yoon & Römer 2020). This dynamic view of academic writing as “social and communicative engagement” (Hyland 2005: 14) highlights its “interactive and rhetorical character”: textual meanings are seen as “socially mediated, influenced by the communities to which writers and readers belong” (Hyland 2004: 12). Mastering these rhetorical strategies and the linguistic means whereby they can be conveyed is therefore crucial for novice academic writers as they enter the academic community.

This paper explores the ways novice and expert academic writers employ hedges, a category of the interactional dimension of metadiscourse, defined by Hyland (2005: 52) as devices

which indicate the writer's decision to recognize alternative voices and viewpoints and so withhold complete commitment to a proposition. Hedges emphasize the subjectivity of a position by allowing information to be presented as an opinion rather than a fact and therefore open that position to negotiation.

Hedges can therefore be considered markers of epistemic stance, i.e. “meanings of certainty, doubt, actuality, precision or limitation, as well as ... the source or perspective of knowledge” (Gray & Biber 2012: 17).

Novice academic writers have been shown to differ from expert writers to a large extent in “conveying their attitudes and critical perspectives” (Yoon & Römer 2020: 210). The variation may be due not only to language proficiency and different levels of awareness of academic writing conventions, but also to the differences in the genres the two groups of writers produce “as students tend to write in response to curricular assignment genres rather than professional genres such as research articles” (Aull 2019: 268). Nevertheless, in the grading of students' academic assignments, “use of stance conventions directly influences scores received by native speaker and English learner student writers alike” (ibid.).

While the use of hedging strategies and devices is “generally acknowledged to be difficult for learners to acquire” (Hyland 2005: 133), this may be further complicated for non-native English novice academic writers by intercultural differences in expressing stance (Dahl 2004). Czech academic discourse has traditionally been characterised by avoidance of tension, backgrounded authorial presence and low level of interaction with the reader. This is reflected, for instance, in the frequent use of impersonal structures or in the fact that “authors' statements of claims and purposes are often hedged by modals and lexical verbs expressing tentativeness and humility” (Dontcheva-Navratilova 2015: 54).

2 Phraseology and learner corpora

Kjellmer's (1991: 12) observation that language communication relies, to a large extent, on “combinations of words that customarily co-occur” has been frequently cited over the past thirty years in connection with the frequency-based approach to phraseology, stressing the fact that “the language we use every day is composed of prefabricated expressions, rather than being strictly compositional” (Gray & Biber 2015: 125). This approach, informed by corpus studies, has shown that “meaning is attached to frequently

occurring sequences rather than to their constituent lexical or grammatical items” (Hunston 2003: 32; cf. e.g. Sinclair 1991, Groom 2005). These sequences can be seen as constituting the phraseology of a language. From the functional point of view, phraseology can be understood as “the preferred way of saying things in a particular discourse” (Gledhill 2000: 1). Therefore, “[r]ecurrent word combinations do not only contribute to idiomaticity, but also contribute to demonstrating membership in a specific discourse community” (Ädel & Erman 2012: 81). We may speak of the phraseology of a register or genre, since

different written and spoken genres and different discourse communities... select or prioritise different phraseological patterns; the former on the grounds that they serve different communicative and institutional purposes and thus prioritise different rhetorical strategies... and the latter on the grounds that they are characterised by different ideational interests and interpersonal practices (Groom 2005: 258).

For novice academics, the awareness and appropriate choice of genre-specific phraseology is not only a key to comprehension and fluency reducing the processing effort (Nesselhauf 2005), but also serves as “badge of identity”, reflecting their “desire to behave like those with whom ... [they] most closely identify, or who embody ... [their] aspirations” (Wray 2005: 595). Phraseology is “one of the aspects that unmistakably distinguishes native speakers of a language from L2 learners” (Granger & Bestgen 2014: 229). At the same time,

[t]he fact that native and non-native apprentice academic writers lack very similar sets of expert academic English phraseological items in their papers indicates that both groups of students may need similar training or help with their academic writing on their way to becoming more proficient writers (Römer 2009: 99, cf. Dontcheva-Navratilova 2012: 55-56).

Research into phraseology is closely tied to learner corpus research, “since corpus analysis lends itself especially well to the study of recurrent multi-word units” (Ebeling & Hasselgård 2015: 208).

The present paper argues that when exploring the use of hedges by novice (English L1 and L2) and English L1 expert academic writers a phraseological perspective can be adopted, focussing on multi-word units which convey the writers’ epistemic stance. Its research goals are therefore twofold: a) test inductive corpus-driven methods as an approach to the identification of hedging patterns in academic texts; b) compare the hedging patterns used by the three groups of academic writers; the fact that the groups differ along two dimensions – language proficiency and academic experience – may highlight the two types of challenge that novice English L2 writers have to face, viz. the linguistic challenge of English as a foreign language, and the academic challenge of entering the academic discourse community (Malá 2020).

3 Corpora used

The approach presented here combines phraseology and learner corpus research to explore the use of hedges by three groups of academic writers – English L2 novice writers (Czech university students of the BA English Studies programme), English L1 novice writers (students’ essays from the BAWE corpus¹, compiled at British universities), and English L1 professional writers (represented by articles published in academic journals). The characteristics of the three corpora are summarized in Table 1. I tried to reduce the impact of subject matter as much as possible since discipline seems to be a strong cause of differences (Ebeling & Hasselgård 2015). The three corpora, however, differ in size, with the L2-novice corpus being the smallest in terms of tokens, but comparable to the L1-novice corpus in the number of texts, and the L1 corpora comprising similar numbers of tokens, but differing in the number of texts. Sketch Engine² was used as a tool for both compilation and analysis of the corpora.

corpus	L2-novice	L1-novice	L1-pro
L1	Czech	English	English
academic level	novice	novice	expert
source	Charles University, Prague, Faculty of Arts, English Studies Programme, BA	BAWE, universities of Warwick, Reading and Oxford Brookes, English	academic journals
time	2016–2019	2004–2017	2000–2015
register	students’ essays	students’ essays	academic papers
topic	English literature: Renaissance to Restoration	English literature	English literature: Renaissance
size: words / tokens (approx.)	127,400 / 145,500	228,200 / 273,300	256,800 / 307,000
No. of texts	80	93	34

Table 1: The corpora used in the present study

4 The identification of hedging patterns

One of the limitations of the study of metadiscourse, as pointed out by Hyland (2005: 58), consists in the identification of metadiscourse devices: “... metadiscourse studies deal

only with *explicit* devices which can be clearly identified in the text”, such as the hedges *would, may, could, might, seem, assume, suggest, possible, perhaps, about, or possibly*.

Since previous studies have shown convincingly that meaning resides in multi-word units rather than in their constituent lexical or grammatical items (see Section 2), in the present study, I approach interactional metadiscourse from the phraseological perspective, focussing on recurrent patterns,

which are combinations of lexis and grammar, and which typically consist of a partly fixed lexical core plus other variable items. The form of these constructions can be specified as combinations of collocation and colligation (co-occurring words and grammatical categories) and their meanings can be specified as often evaluative communicative functions (Stubbs 2007: 164).

The method employed here to identify the phraseology pertaining to the expression of epistemic stance can be described as corpus-driven or inductive: corpus data and search software is used “in order to automatically extract patterns which are invisible without such help” (Stubbs 2007: 164). I then try to analyse the functions of these patterns (taking contextual cues into account), and make generalizations about them.

Two similar approaches to the identification of hedging patterns are adopted here. Both take a single word as its starting point. 3-grams, i.e. recurrent contiguous sequences of three words, which comprise the ‘core’ word are then extracted automatically from the three corpora. Functionally similar 3-grams which display a certain degree of lexical variation can be lumped together, constituting a pattern. The term ‘pattern’ is used here to refer to a recurrent multi-word lexico-grammatical unit (allowing some internal variation) which performs a specific function in the text.

The two approaches I use in the two studies described in Sections 5 and 6 differ in what constitutes the ‘core’ item in the initial step. In the first, more corpus-driven study, the ‘core’ item in the phraseological sequences is the pronoun *it* – one of the most frequent grammatical words overused by novice academic writers. Although they carry no evaluative meaning themselves, “small” grammatical words were shown to be “crucial to textual meaning” in specialised corpora (Hunston 2008: 272), “because they reveal phraseologies that are linked to recurrent meanings and functions rather than to subject-matter” (ibid.: 292).³

The second study takes the verb *seem*, itself an epistemic stance marker, as the starting point for the identification of recurrent phraseologies used to imply that “a statement is based on the writer’s interpretation rather than certain knowledge” (Hyland 2005: 68).

5 Study 1: *it*-patterns

As an initial step towards the hedging patterns I identified the grammatical words (see Section 4) which are significantly⁴ over- or underused by novice academic writers, with the English L1 published papers (L1 pro) serving as the norm (Table 2). The word-forms *is*, *it*, *be*, *can* are significantly overused by both groups of novice writers, *of* and *that* are underused by novice writers in comparison with English L1 expert writers.

frequency per million tokens	L2 novice	L1 novice	L1 pro
<i>is</i>	21,019.4	17,121.7	11,038.8
<i>it</i>	9,093.7	6,337.9	5,263.7
<i>be</i>	6,358.0	4,317.9	3,622.0
<i>can</i>	3,216.8	2,495.6	1,635.1
<i>of</i>	3,612.6	33,749.4	39,523.3
<i>that</i>	10,420.3	10,487.5	12,090.9

Table 2: Grammatical words overused (dark grey) or underused (light grey) by novice academic writers in comparison with L1 expert writers

In the present paper I will focus on the phraseology surrounding the pronoun *it*. The pronoun was selected not only due to its frequency but also because it has been attested in patterns whose occurrence in academic texts is “motivated by genre-specific purposes and discipline-specific practices” (Groom 2005: 258). The most frequent 3-grams comprising the word *it* extracted from the two novice corpora are listed in Table 3.⁵

	L2 novice		L1 novice		
	frequency (raw)	freq. per million tokens		frequency (raw)	freq. per million tokens
<i>it is not</i>	43	295.6	<i>that it is</i>	49	179.3
<i>it is the</i>	38	261.2	<i>it is the</i>	42	153.7
<i>it is a</i>	32	220.0	<i>it is possible (to/that)</i>	42	53.7
<i>it can be</i>	27	185.6	<i>it is a</i>	38	139.1
<i>that it is</i>	21	144.3	<i>as it is</i>	36	131.7
<i>but it is</i>	19	130.6	<i>it is not</i>	33	120.8
<i>and it is</i>	17	116.9	<i>it could be</i>	27	98.8
<i>as it is</i>	16	110.0	<i>it is clear</i>	19	69.5
<i>it does not</i>	14	96.2	<i>it might be</i>	17	62.2
<i>it is also</i>	13	89.4	<i>it can be</i>	17	62.2
<i>it might be</i>	11	75.6	<i>and it is</i>	16	58.6
<i>it is clear</i>	11	75.6	<i>it is interesting</i>	15	54.9
<i>it seems that</i>	11	75.6	<i>but it is</i>	14	51.2

Table 3: The most frequent 3-grams comprising the word *it* extracted from the two novice corpora

The patterns with modal and epistemic verbs and adjectives, which may potentially be involved in hedging, are of particular interest here. Czech students, generally, use the modal

verbs *can*, *might* and *could* more frequently than expert academic writers, and underuse *may*.⁶ The most prominent difference, however, rests in the patterns of use of these modal verbs. Czech students prefer the pattern with initial *it* and an extraposed *that*-clause (Example 1, 26 instances). The repertoire of verbs employed in the pattern is limited to verbs of speaking or argumentation. The pattern can be schematically represented as ‘*it can/might/could/may be V-ed (that) + clause*’. The verb forms used in the pattern comprise *argued*, *seen*, *said*, *assumed*, *noticed*, *understood*, *proposed*, *pointed out*, *shown*, *observed*, *speculated*. There is a particularly strong association between the verb *argue* and the *it*-pattern with modal verbs: out of the 46 instances of the verb in the L2 corpus, 23 per cent occur in this pattern.

- (1) *Even though, the author criticized her, **it can be argued** she is also a victim of certain expectations from society, ...* (L2-novice)

This pattern serves as a hedge, making it possible to report “the attitude of the speaker or writer of the text, even though the author does not assume explicit responsibility for the attitude” (Biber et al. 1999: 661). “[The] writer uses a variety of modality markers to give an indication of the degree of probability, value or necessity of content”, indicating at the same time “the non-factual status of a proposition by marking it as being their suggestion, contention, argument, assumption, and so on” (Hewings & Hewings 2014: 105).

In professional academic writing, the uses of the modal verbs as well as the lexical verbs they collocate with are more diverse (Example 2), with a single occurrence of the above initial *it* pattern attested in the corpus. This is in line with Ädel and Erman’s (2012: 87) observation that “native speakers...draw on a wider variety of lexical resources for expressing uncertainty and doubt”.

- (2) *In this context **it may be worth noting** that Ate or Mischief is the name that Jonson’s masque gives to the witches’ leader.* (L1-pro)

Another pattern overused by L2 writers – *it may be that* – is illustrated in Example (3). Due to its similarity to the corresponding Czech pattern *je možné, že*, we may assume a transfer effect here. One instance of the pattern, however, is attested in an L1 published paper too.

- (3) ***It may be that** Jessica is somewhat angry about the fact that she needs to disguise herself as a man ...* (L2-novice)

6 Study 2: *seem*-patterns

This section explores the phraseology surrounding the verb *seem*. *Seem* is an epistemic evidential marker (Dontcheva-Navratilova 2018); the patterns including the verb are therefore likely to express modality and writers' stance.

The collocations and colligations of the verb *seem* revealed differences between the patterns dominant in L2 novice and L1 expert texts, with the language choices reflecting the rhetorical strategies employed by the two types of academic writers.

Starting from quantitative findings, Table 4 shows the frequencies of the lemma *seem* in the three corpora. The results indicate significant overuse of the verb by L2 novice writers (compared to L1 expert writers), and the frequency in L1 novice texts in between L2 and L1 pro. This itself may suggest a scale of subjectivity and commitment to a proposition with L1 expert writers and L2 novice writers representing opposite strategies.

	L2 novice		L1 novice		L1 pro	
	Freq. raw	Freq. relative	Freq. raw	Freq. relative	Freq. raw	Freq. relative
<i>seem</i>	149	1,024.2	250	914.8	243	791.5

Table 4: Frequency of the lemma *seem* (relative frequencies are calculated per million tokens)

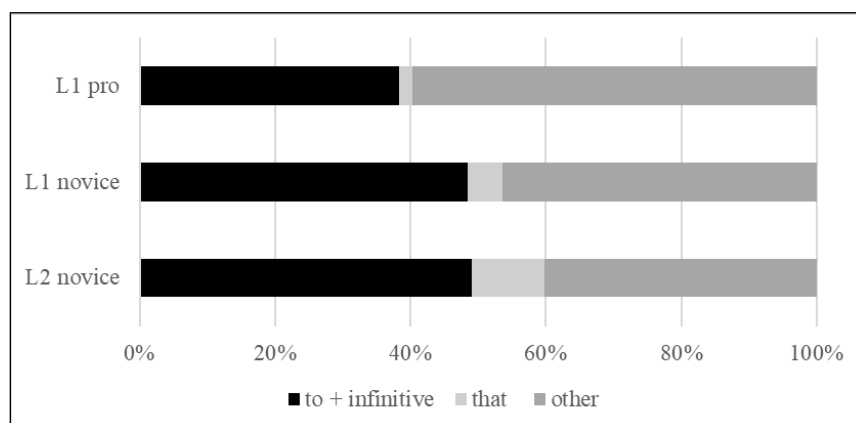


Figure 1: Complementation of the lemma *SEEM*

At the same time, there are differences among the three groups of writers in the patterns associated with the verb. Focussing on colligation, we can observe a difference in the representation of complementation forms among the three corpora (Figure 1). While L2 and L1 novice writers rely on complementation by the infinitival clause in 49.0 and 48.4 per cent, respectively, of their uses of the lemma *seem*, expert L1 writers display considerably more variation in the complementation patterns. In published articles, *seem* is often complemented

by an adjective phrase in constructions with anticipatory *it* and an extraposed clausal subject, finite or infinitival (Example 4a). Czech students, on the other hand, appear to be inclined to use the structurally simpler construction with *it seems that* (Example 4b, see also Figure 2). The effect of using the verb *seem* is “to suggest indecision about the statement made in the *that*-clause” (Hewings & Hewings 2014: 108). The pattern thus functions as a hedge; with the *it* subject adding the sense of detachment and lack of commitment. This pattern is often preceded by causal or concessive conjunctions or adverbials (*therefore, so, even if*), indicating that it is used primarily to draw tentative conclusions or propose alternative interpretations.

- (4) a. *Yet it seems highly unlikely that Volpone is simply polemicizing against Catholicism...* (L1-pro)
 b. *So even if it seems that everything in life comes to an end, there are things that surpass or have the power to overcome the obstacles of Time.* (L2-novice)

To explore the phraseology associated with *seem* in more detail, I compared the 3-grams comprising the lemma *SEEM* in the three corpora. Figure 2 shows the 3-grams shared by all groups of writers.⁷ As we shall see below, the differences among the three groups of writers again do not rest solely in the frequency, but rather in the specific functions of the patterns.

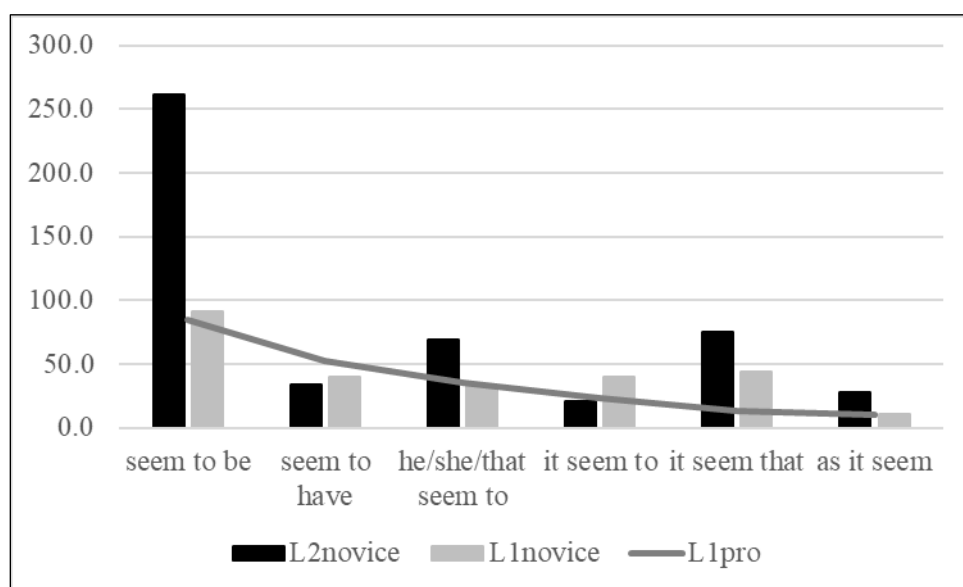


Figure 2: Patterns including the lemma *seem* attested in all three corpora

SEEM to be

What appears to be the same pattern in the three corpora, overused by L2 novice writers, comprises two major patterns with a different verb *be*. In L1 writers’ texts, both novice and professional, this is often a pattern with a progressive infinitive (40% instances in

L1-novice, 27% in L1-pro corpus; Example 5a). Where *be* is a copula in the L1 corpora, it is typically complemented by a classifying noun phrase (Example 5b). For Czech writers, complex non-finite forms are generally more difficult to use due to the narrow repertoire and low frequency of corresponding forms in their mother tongue (Dušková et al. 2009: Chapter 15). The pattern *seem to be* is most often followed by an adjective phrase in the L2 novice corpus (47% instances of the pattern, Example 5c). The progressive forms are few (10.5%).

- (5) a. *In establishing enumerated similarities between himself and Hester, Hawthorne **seems to be trying** to validate his ability to assume the female voice.* (L1-novice)
 b. *As the growth of spiritual idleness, **Orgoglio would seem to be a materialization** of that Augustinian ailment which fells Verdant...* (L1-pro)
 c. *This effort, presented by Gambino, **seems to be impossible**.* (L2-novice)

SEEM to have

All three groups of writers use the pattern predominantly with the perfect infinitive (Example 6a). In novice academic texts, both L1 and L2, *have* may also appear as a lexical verb (36 and 20% instances, respectively; Example 6b). However, the frequencies are too low to allow any generalization.

- (6) a. *...so few modern critics **seem to have** noticed this.* (L1-pro)
 b. *Instead it **seems to have** the opposite effect.* (L1-novice)

it SEEM to, he/she/that SEEM to

Apart from the subject raising constructions, this pattern also comprises instances with a final prepositional phrase, indicating explicitly the source of the epistemic evaluation. It is only in the expert corpus though that the stance is ascribed to the writer (*to me*; Example 7a); in the novice corpora, the pattern with the prepositional phrase is not a hedge (Example 7b).

- (7) a. *In fact **it seems to me** that neither worked from a manuscript at all.* (L1-pro)
 b. *he realized nothing is as **it seemed to him** before* (L2-novice)

as it SEEM

The overuse of the pattern by Czech students can probably be accounted for by the corresponding Czech pattern *jak se zdá*, used in clause initial or medial position (Example 8a). L1 writers, on the other hand, mostly use the pattern as a causal one, without a hedging effect.

A functionally similar *seem*-pattern used by Czech students – *it seems as if* – was not attested in the L1 corpora. Expert writers, on the other hand, employed another parenthetical hedging *seem*-pattern – , *it seems*, – most often placed in medial position in the clause,

following the subject (Example 8b). L1 novice writers use the pattern too, albeit less frequently.

- (8) a. *As it seems the author is upset with the society because it keeps believing politicians and their orations.* (L2-novice)
b. *The comic taxonomy, it seems, is simply not sufficient, especially given assessments such as Norman Rabkin's...* (L1-pro)

Another hedging pattern surrounding the lemma *seem* – *it seems likely that* – was attested only in the L1 expert writers' articles (Example 9). While the pattern “may appear on the surface to be impersonal and objective, in fact the adjective choice opens up a space which the writer can use to indicate the nature of their comment on what follows (Charles 2004: 78).

- (9) *It also seems likely that a man named Smith might have been less willing to lend Elizabeth seven pounds...* (L1-pro)

The hedging *seem*-patterns were found to comprise epistemic modal verbs in all three corpora, enhancing tentativeness. There, however, seems to be a difference between the patterns used by expert writers and those attested in novice texts. In the L1-pro corpus the pattern comprises a finite or infinitival clause – ‘*might/may/would seem to-/that* clause’ – with subject raising. The pattern is oriented cataphorically, relativizing the content of the dependent clause (Example 10a). Even though novice writers rely on the same combination of a modal verb and *seem*, they were found to prefer a pattern with an adjective or noun phrase – ‘*might/may/would seem* AdjP/NP’, conveying a tentative evaluation of the preceding statement. L2 novice writers often use the pattern comprising a modal verb and *seem* to introduce an interpretation in a non-committed way, only to refute the claim in the following clause (Example 10b).

- (10) a. *By mirroring Caesar's magnanimity, Antony and Octavius might seem to be attempting to usher in a new period of calm and peace.* (L1-pro)
b. *This may seem incompatible with the theory of mutability - that everything changes, but it is not.* (L2-novice)

7 Conclusion

The analysis of the recurrent patterns expressing the writer's stance has highlighted the two types of challenge English L2 novice academic writers have to face when writing their early academic texts, namely the ‘academic’ and the ‘linguistic’ challenge, which are particularly prominent in the sphere of interactional metadiscourse. For university students, both English L2 and L1, it is a challenging task to strike the balance between signalling credibility and expressing their claims with due caution. What has been termed ‘academic’

challenge here mirrors their lack of experience in expressing their claims. The impact of the ‘academic’ challenge is manifested in the overuse of grammatically relatively simple hedging patterns with little lexical variation, such as ‘*it seems that* + clause’ or ‘*it can/might/could/may be* V-ed (*that*) + clause’. In these patterns the evaluative expressions (*seem*, modal verbs) are separated from the statement made in the dependent clause, which contributes to their broad applicability. The impersonal subject *it* further allows the writers to distance themselves from the proposition. A similar separation of the evaluated proposition from the evaluation itself can be seen in the patterns with anaphoric subjects followed by a copular predication with a modal verb ‘*might/may/would seem* AdjP/NP’. Czech students were found to use this pattern to acknowledge the possibility of other interpretations before giving their own opinion. What novice academic essays lack is grammatically more complex hedging patterns with subject raising and complex verb forms, and explicit attribution of stance to the writer using patterns with prepositional phrases (*it seems to me*). As pointed out by previous studies, the range of epistemic stance patterns employed by L2 writers is smaller than that manifested in L1 texts (cf. Chen & Baker 2010: 41).

Despite their high level of proficiency in English, Czech novice academic writers have to face linguistic challenges, some ascribable to transfer, some related more generally to the typological differences between English and Czech. A manifestation of the former can be seen in the L2 writers’ use of the pattern *as it seems*, corresponding to the parenthetical Czech pattern *jak se zdá*. The latter may account for the reluctance of Czech students to use complex non-finite verb forms non-existent in Czech.

The scope of the present study is quite limited, but I hope to have demonstrated the possibilities of approaching hedges, and perhaps metadiscourse more generally, from a phraseological perspective. The expression of epistemic stance was shown to rest not on individual words but rather on multi-word lexico-grammatical sequences, allowing for some degree of lexical variation. The method adopted here, drawing on recurrent multi-word patterns, may provide some answers to the questions of identification of metadiscourse devices in academic texts. While the initial steps in exploring hedging patterns were corpus-driven, starting either from grammatical words or the epistemic verb *seem*, the recurrent patterns were analysed from the functional point of view. In this way, differences between English L2 and L1 writers on the one hand, and those between novice and expert ones on the other, were found to consist not only in over- or underuse of a pattern but also in the ways it is employed as a hedging device by each group of writers. This may also have teaching

implications – English L2 novice academic writers should be made aware not only of the lexical means available to them for the expression of epistemic stance (such as the verbs *seem* or *suggest*), but also of the patterns that convey the interpersonal meanings (such as ‘*it seems* AdjP to V-infinitive’). At the same time, the results of the present study suggest that even at a high level of proficiency, attention should be paid to areas of the English language that Czech students may find difficult. Awareness of language- and register-specific means of expressing stance should make it easier for English L2 novice academic writers to “adapt their rhetorical strategies as they enter a new discourse community” (Yoon & Römer 2020: 209).

Notes

¹ BAWE: <https://www.coventry.ac.uk/research/research-directories/current-projects/2015/british-academic-written-english-corpus-bawe/>

² Sketch Engine: <https://www.sketchengine.eu/>

³ For a similar approach, exploring the phraseological patterning around five grammatical keywords (*of, and, that, as* and *this*) cf. Whiteside & Wharton (2019).

⁴ Significance was assessed using *Significance and Effect Size Calculator*, (available at <http://ucrel.lancs.ac.uk/llwizard.html>) with the log-likelihood value of 3.8 or higher being significant at $p < 0.05$.

⁵ In the L1-pro corpus, the *it* 3-grams of comparable relative frequency and rank did not include any modal expressions. This may be connected with the greater variability of hedging patterns in expert L1 writing.

⁶ The relative frequencies (per million tokens) of the modal verbs *can, might, could* and *may* in the L2 / L1 pro corpora are: *can* 3,223.7 / 1,680.7; *might* 797.3 / 550.5; *could* 742.3 / 719.8; *may* 604.9 / 918.5.

⁷ The cut-off point was set at 2 instances at least in the L2 corpus, and 3 in the L1 corpora. Three 3-grams with anaphoric pronouns in initial position were merged in Figure 2 (*he/she/that seem to*). The 3-grams are lemmatized.

References

- Abdi, R. (2002) ‘Interpersonal metadiscourse: An indicator of interaction and identity.’ *Discourse Studies* 4(2), 139–145. <https://doi.org/10.1177/14614456020040020101>
- Ädel, A. and Erman, B. (2012) ‘Recurrent word combinations in academic writing by native speakers and non-native speakers of English. A lexical bundles approach.’ *English for Specific Purposes* 31(2), 81–92. <https://doi.org/10.1016/j.esp.2011.08.004>
- Aull, L. (2019) ‘Linguistic markers of stance and genre in upper-level student writing.’ *Written Communication* 36(2), 267–295. <https://doi.org/10.1177/0741088318819472>
- Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. (1999) *Longman Grammar of Spoken and Written English*. Edinburgh: Pearson Education Ltd.
- Bruce, I. (2016) ‘Constructing critical stance in university essays in English literature and sociology.’ *English for Specific Purposes* 42, 13–25. <https://doi.org/10.1016/j.esp.2015.10.005>
- Charles, M. (2004) *The Construction of Stance. A Corpus-based Investigation of Two Contrasting Disciplines*. Unpublished doctoral dissertation. University of Birmingham.
- Chen, Y. and Baker, P. (2010) ‘Lexical bundles in L1 and L2 academic writing.’ *Language Learning & Technology* 14(2), 30–49.
- Dahl, T. (2004) ‘Textual metadiscourse in research articles. A marker of national culture or of academic discipline?’ *Journal of Pragmatics* 36(10), 1807–1825. <https://doi.org/10.1016/j.pragma.2004.05.004>
- Dontcheva-Navratilova, O. (2012) ‘Lexical bundles in academic texts by non-native speakers.’ *Brno Studies in English* 38, 37–58. <https://doi.org/10.5817/BSE2012-2-3>
- Dontcheva-Navratilova, O. (2015) ‘The changing face of Czech academic discourse.’ In: Bennett, K. (ed.) *The Semiperiphery of Academic Writing*. Basingstoke: Palgrave Macmillan. 39–61. https://doi.org/10.1057/9781137351197_3

- Dontcheva-Navratilova, O. (2018) 'Intercultural and interdisciplinary variation in the use of epistemic lexical verbs in linguistics and economics research articles.' *Linguistica Pragensia* 28(2), 154–167.
- Dušková, L., Knittlová, D., Peprník, J., Strnadová, Z. and Tárníková, J. (2009) *Mluvnice současné angličtiny na pozadí češtiny*. Electronic edition. Ústav anglického jazyka a didaktiky, FF UK. 1 January 2022 <http://www.mluvniceanglictiny.cz>
- Ebeling, S. and Hasselgård, H. (2015) 'Learner corpora and phraseology.' In: Granger, S., Gilquin, G. and Meunier, F. (eds) *The Cambridge Handbook of Learner Corpus Research*. Cambridge: Cambridge University Press. 207–230. <https://doi.org/10.1017/CBO9781139649414.010>
- Gledhill, C. (2000) *Collocations in Science Writing*. Tübingen: Gunter Narr.
- Granger, S. and Bestgen, Y. (2014) 'The use of collocations by intermediate vs. advanced non-native writers. A bigram-based study.' *IRAL - International Review of Applied Linguistics in Language Teaching* 52(3), 229–252. <https://doi.org/10.1515/iral-2014-0011>
- Gray, B. and Biber, D. (2012) 'Current conceptions of stance.' In: Hyland, K. and Sancho Guinda, C. (eds) *Stance and Voice in Written Academic Genres*. Basingstoke: Palgrave Macmillan. 15–33. https://doi.org/10.1057/9781137030825_2
- Gray, B. and Biber, D. (2015) 'Phraseology.' In: Biber, D. and Reppen, R. (eds) *The Cambridge Handbook of English Corpus Linguistics*. Cambridge: Cambridge University Press. 125–145. <https://doi.org/10.1017/CBO9781139764377.008>
- Groom, N. (2005) 'Pattern and meaning across genres and disciplines. An exploratory study.' *Journal of English for Academic Purposes* 4(3), 257–277. <https://doi.org/10.1016/j.jeap.2005.03.002>
- Hewings, A. and Hewings, M. (2014, 1st ed. 2004) 'Impersonalizing stance. A study of anticipatory *it* in student and published academic writing.' In: Coffin, C., Hewings, A. and O'Halloran, K. (eds) *Applying English Grammar. Functional and Corpus Approaches*. London and New York: Routledge. 101–116.
- Hunston, S. (2003) 'Lexis, wordform and complementation pattern. A corpus study.' *Functions of Language* 10, 31–60. <https://doi.org/10.1075/fof.10.1.03hun>
- Hunston, S. (2008) 'Starting with the small words. Patterns, lexis and semantic sequences.' *International Journal of Corpus Linguistics* 13(3), 271–295. <https://doi.org/10.1075/ijcl.13.3.03hun>
- Hyland, K. (2004) *Disciplinary Discourses*. Ann Arbor: The University of Michigan Press.
- Hyland, K. (2005) *Metadiscourse. Exploring Interaction in Writing*. London and New York: Continuum.
- Kjellmer, G. (1991) 'A mint of phrases.' In: Aijmer, K. and Altenberg, B. (eds) *English Corpus Linguistics. Studies in Honour of Jan Svartvik*. London and New York: Longman. 111–127.
- Lancaster, Z. (2016) 'Expressing stance in undergraduate writing: Discipline-specific and general qualities.' *Journal of English for Academic Purposes* 23, 16–30. <https://doi.org/10.1016/j.jeap.2016.05.006>
- Malá, M. (2020) 'Phraseology in learner academic English. Corpus-driven approaches.' *Discourse and Interaction* 13(2), 75–88. <https://doi.org/10.5817/DI2020-2-75>
- Nesselhauf, N. (2005) *Collocations in a Learner Corpus*. Amsterdam: John Benjamins. <https://doi.org/10.1075/scl.14>
- Römer, U. (2009) 'English in academia. Does nativeness matter?' *Anglistik: International Journal of English Studies* 20(2), 89–100.
- Sinclair, J. M. (1991) *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Stubbs, M. (2007) 'Quantitative data on multi-word sequences in English: The case of the word world.' In: Hoey, M., Mahlberg, M., Stubbs, M. and Teubert, W. (eds) *Text, Discourse and Corpora*. London and New York: Continuum. 163–190.
- Whiteside, K. L. and Wharton, S. (2019) 'Semantic patterning of grammatical keywords in undergraduate academic writing from two close disciplines.' *Journal of English for Academic Purposes* 39, 1–20. <https://doi.org/10.1016/j.jeap.2019.03.005>
- Wray, A. (2005) 'Formulaic language.' In: Brown, K. (ed.) *Encyclopedia of Language and Linguistics*. Oxford: Elsevier. 590–597. <https://doi.org/10.1016/B0-08-044854-2/04777-5>

Yoon, H. and Römer, U. (2020) 'Quantifying disciplinary voices. An automated approach to interactional metadiscourse in successful student writing.' *Written Communication* 37(2), 208–244. <https://doi.org/10.1177/0741088319898672>

Acknowledgement

This work was supported by the Czech Science Foundation grant 19-05180S Phraseology in English academic texts written by Czech advanced learners: a comparative study of learner and native speaker discourse.

Markéta Malá is Associate Professor of English language and linguistics at the Faculty of Education, Charles University, Prague. The main focus of her work is on contrastive corpus-assisted approaches to phraseology, academic English and the language of children's literature.

Address: Markéta Malá, Department of English Language and Literature, Faculty of Education, Charles University, Magdalény Rettigové 4, 116 39 Praha 1, Czech Republic. [e-mail: marketa.mala@pedf.cuni.cz]

A COMPARISON OF CLIMATE CHANGE LEXICAL CREATIVITY AMONG AMERICAN, EUROPEAN AND INTERNATIONAL TWITTER USERS

Vanessa Marcella

<https://doi.org/10.5817/CZ.MUNI.P280-0212-2022-6>

Abstract

Climate change is an environmental, social, cultural and political phenomenon which has led to a bitter controversy among political leaders, social movements, and activists. The aim of this case study is to analyze lexical innovations in climate change discourse in the micro-blogging service Twitter, and more in particular, in the use of English language through hashtags by a new generation of young activists, NGOs, and American and European politicians and organizations in the time frame from 2015 to 2020.

By means of this case study, we prove that, through the creativity of hashtags, Twitter users can take part in the heart of a discussion related to climate change leading to a valuable interpretive framework.

Keywords

climate change, lexical creativity, corpus linguistics, Twitter, hashtags

1 Introduction

As a global phenomenon, climate change is discussed by all types of population by any kind of means of communication. This has led to one of the greatest divergences about climate change, which lies in the multitude and contrasting messages we receive and the different ways we interpret them (Hulme 2009). In literature critical discourse analysis, framing theories, and cognitive linguistics have been the main approaches used to identify underlying ideologies (Killingsworth & Palmer 1992, Halliday 2001). Past studies primarily focused on the discourse of environmental politics (Hajer & Versteeg 2005). More recently, studies have analyzed climate change in social-networking-based communications, investigating linguistic variations and lexical creativity in environmental discourses (Nerlich & Koteyko 2009, Nerlich al. 2010, Koteyko & Atanasova 2016). For instance, Segerberg and Bennett (2011) studied climate change protests related to the 2009 15th United Nations Conference of the Parties on Climate Change (COP-15) in Copenhagen by analyzing random samples of tweets related to the event. Bruns and Burgess (2011) analyzed the public

discourse, including climate change, during the 2010 Australian elections. Kirilenko and Stepchenkova (2014) investigated the global public discourse on climate change Twitter in five main languages from 2012 to 2013 based on the collection of 1.8 million tweets that included the words “climate change” and/or “global warming”. Veltri and Atanasova (2015) carried out a study exploiting different features deriving from Twitter data: they collected a random weekly sample of tweets in March 2013 containing the keywords “climate change” and “global warming” and carried out text mining and semantic networks, and a psychological process-based classification of tweets. Finally, Jang and Hart (2016) published a study on frames and terms used by tweeters about “climate change” and “global warming” in four English speaking countries retrieving tweets over two years. Most of these studies dedicated to lexical variation in social networks (henceforth SN) focus either on synchronic or local phenomena. Therefore, we propose a comparative, long-term study, on climate change as a global and enduring discussion, involving international stakeholders, particularly focusing on lexical variation in discourse use in relation to climate change on SN Twitter.

2 Hashtags and lexical creativity

Hashtags have been taken into consideration due to their multi-functional linguistic features in online discourse over time. They can be considered as keywords which allow users to follow a specific topic or might be used as search terms contributing to the creation of a virtual community which shares a particular interest. Zappavigna (2011) defines hashtags as a typographic convention used to mark the topic of a tweet, they are a kind of “inline” metadata integrated into the linguistic structure of the tweets which usually explicitly refer to the topic of the tweet (Zappavigna 2012, 2015). In fact, the hashtag in particular enables people to connect online, and thanks to the easiness of its use, anyone can create a new one, and therefore, create new words. In turn, all users can use and modify a hashtag increasing the spread of its usage, and anyone can see whether a hashtag is used by other users or not (Pizarro Pedraza & De Cock 2018). The main characteristic of a hashtag is the ability to increase the “loudness” of a discourse related to a specific topic, making it viral and searchable, and also creating a more solid bond within the virtual community (Zappavigna 2011, Roginsky & De Cock 2015, Schwell 2015). When this type of bond is established among a large number of Twitter uses, and in this case through the use of hashtags which contain the string *clim**, the so called “ambient affiliation” is formed (Bruns & Burgess 2011,

Zappavigna 2011). Specifically, as SN sites and social media become pervasive, the affordances and dynamics of networked publics can show why people engage the way they do (Boyd 2010).

While from a linguistic point of view, hashtags might appear as neologisms in relation to what Crystal (2001) describes as the creation of new lexical items that become part of the language of a determined speech community during a particular period of time. With reference to time, the awareness of the lack of words to express new climate related perceptions, feelings and scenarios, and thus the need to create new ones, derives from what Mühlhäusler (2003) defines as a source of environmental awareness of lexical gaps.

Lexical creativity and neologisms related to the environment, which mostly appear in SN, might have a significant impact on how people perceive climate change in particular, and the environment in general. According to *Webster's New World Dictionary*, a neologism is: a new word or a new meaning for an established word; the use of, or the practice of creating, new words or new meanings for established words. As argued by Rey (1988), neology may be defined as an evolving process which generates new lexical items in a given language and/or in a social community, owing to the emergence of new things, or to the urge to rename things. According to Gerrig and Gibbs (1988) language users may recur to lexical creativity for two main reasons. A first motivation could be the need to “express ideas that are unavailable in the standardized repertory of meanings” (ibid.: 3), intended as a pragmatic reason. A second type of motivation lays on social factors. Thus, since users face new kinds of situation, innovations may occur to achieve new communicative purposes (ibid.). Also, Gupta (1992: 15) considers recurring to creativity as a need to express either an “inner state” or an external goal or problem. This type of lexical creativity may be used to strengthen the sense of connectedness within a community.

Lexical creativity has an important role since it introduces new words and collocations that shed light on the general picture of current human life (Guslyakova et al. 2020). Therefore, the study of neologisms, and more broadly lexical creativity, show the current environmental situation, but might also unveil the reasons that lay behind this climate crisis. They are able to provide solutions or suggest how to live and behave in a more climate-friendly way (ibid.). Creativity is extremely relevant in persuasive discourse where the objective is not only to inform, but also to change opinions (Gerrig & Gibbs 1988: 103). Thus, starting from the assumption that language is a human characteristic, and it can be part of a solution of a problem, it is important to analyze how a general sentiment about climate

change evolves through time, and whether it brings to positive or negative ways of expressing actions, feelings or depicting behavior. The main goal of this study is to analyze the most common hashtags that define climate change as an evolving problem and the least frequent hashtags as an attempt to discover new “climate compounds” throughout time. Accordingly, the two main hypotheses are to investigate lexical creativity firstly assuming that it is more likely to occur in specific groups, and secondly, that it has a negative meaning through time. Since the climate change context is evolving over time, people are more aware of its negative consequences.

3 Data collection and methodology

This study is part of a wider research about lexical creativity on Twitter. Hashtags were retrieved from a Twitter climate-related specialized corpus in English, created for the purpose of this research, in the framework of the Solstice 202CM project¹, following the Cougnon et al.’s (2022) methodology.

Overall, 372,719 tweets were collected from 2015 to 2020 from sixty-three users’ accounts. The first sampling procedure implied the collection of all tweets from specific users, with no time limit. The searched data represent active accounts run by real people. However, all tweets related to former president Trump were downloaded from www.thetrumparchive.com since the @realDonaldTrump account was permanently suspended on 8th January 2021.

These tweets were collected with metadata related to authors’ screen name, time of publication, hashtags, mentions and URLs. For this study, retweets were excluded since they were not relevant to the research aims. A further specialized criterion for this corpus creation refers to place. All tweets were manually annotated taking into account the origin of each actor in order to have a clear distinction of their origin and better understand whether they mostly act in Europe, in the United States or on an international level.

Users’ accounts (Table 1) were divided into specific groups, labeled as “author type”, such as climate activists, NEWS organizations, NGOs, institutional organizations, politicians, and social movements. Specifically, the corpus consists of eight climate activists, four from the US and four from Europe; two news companies, a European and an international one; three NGOs, a European and three international; nine governmental organizations, five from Europe and four from the US; eighteen European politicians and eighteen from the US; and five social movements, one from Europe, two international ones and two from the US.

Authors sorted by author type and origin	Number of tweets
Activists	8,084
EU	4,374
Adélaïde Charlier	104
Anuna De Wever Van Der Heyden	112
Greta Thunberg	1,745
Luisa Neubauer	2,413
US	3,710
Jerome Foster II	3,028
Varshini Prakash	27
Vic J. Barrett	243
Xiye Bastida	412
NEWS	54,577
EU	28,887
Guardian Environment	28,887
INT	25,690
EcoWatch	25,690
NGOs	34,815
EU	6,539
EEB	6,539
INT	28,276
Climate Action Network International	3,522
Greenpeace	24,754
Organizations	43,990
EU	15,582
ENVI Committee Press	1,250
EU Climate Action	2,326
EU Environment	4,086
Greens/EFA	3,844
IPCC	4,076
US	28,408
NOAA	2,628
The Nature Conservancy	10,929
US Department of Energy	9,674
US EPA	5,177

Politicians	151,838
EU	33,148
Adina Valean	409
Andreea Strachinescu	2,621
Antonio Guterres	2,187
Catharina Sikow Magny	1,857
Charles Michel	3,278
David Sassoli	783
Ditte Juul Jorgensen	41
Frans Timmermans	2,818
Jean Lambert	1,483
Kadri Simson	675
Mairead McGuinness	2,809
Margaritis Schinas	1,720
Mariya Gabriel	4,082
Miguel Arias Cañete	1,209
Stella Kyriakides	1,081
Ursula von der Leyen	1,125
Valdis Dombrovskis	4,054
Virginijus Sinkevicius	916
US	118,690
Al Gore	1,246
Alexandria Ocasio-Cortez	7,434
Barack Obama	2,733
Bernie Sanders	12,602
Deb Haaland	6,610
Donald Trump	27,895
Ed Markey	8,753
Andrew Wheeler	560
Gina McCarthy	995
Hillary Clinton	9,285
Jennifer Granholm	9,931
Joe Biden	5,350
John Kerry	1,590
Michael S. Regan	1,734
Mike Pence	4,322

Nancy Pelosi	5,981
Frank Pallone	10,625
Scott Pruitt	1,044
Social movements	79,415
EU	5,964
350.org Europe	5,964
INT	11,274
Extinction Rebellion	10,301
Fridays For Future	973
US	62,177
Climate Reality	53,855
Earth Guardians	8,322
Total	372,719

Table 1: List of Twitter users' accounts and number of tweets related (in raw numbers)

In particular, 94,495 tweets come from European users, 213,144 from Americans, and 65,241 from international ones as illustrated in Figure 1.

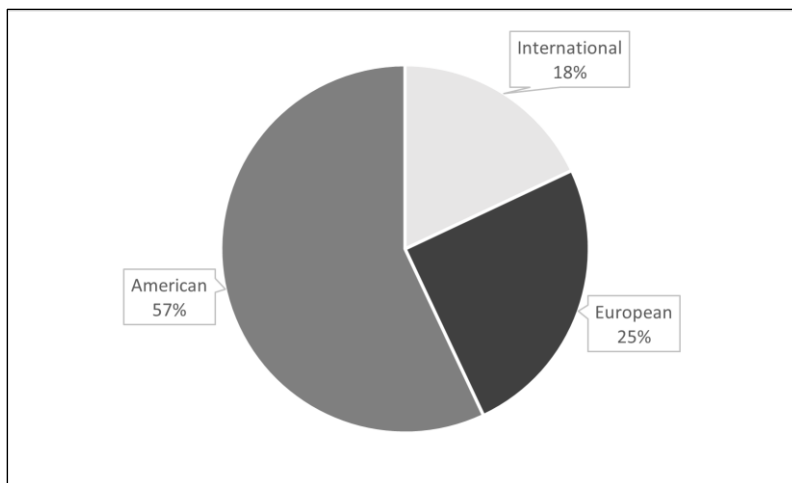


Figure 1: Distribution of tweets by origin in percentages

The second sampling procedure consisted of filtering tweets which contained hashtags. As reported in Table 2, out of 163,753 tweets in English language, 93,502 contain at least one hashtag. Tweets from the US are most likely to include a hashtag, almost the double compared to the European ones, although this group does not include NEWS sites, nor NGOs. European activists are slightly more active than the Americans, while politicians and organizations are more productive in the American context. Tweets from American social

movements are almost nine times more than the Europeans, and more than four times compared to the international ones. However, a great contribution is given by international users (29,371 tweets), and NGOs are the most productive within this group, also compared to the European NGOs.

Stakeholder	EU	INT	US	Total
Activists	793		529	1,322
NEWS	531	9,337		9,868
NGOs	3,733	13,397		17,130
Organizations	8,452		8,793	17,245
Politicians	5,163		5,609	10,772
Social movements	3,161	6,637	27,367	37,165
Total	21,833	29,371	42,298	93,502

Table 2: Number of tweets in English, which contain hashtags, per user type

4 Analysis

Out of the 93,502 tweets containing hashtags, a whole set of 861 hashtags was collected from which 717 hashtags containing the string *clim** (i.e. all words beginning with *clim*) were identified. Table 3 presents the ten most frequent hashtags starting with the string *clim**.

	Hashtag	Frequency
1	climatechange	9,195
2	climatehope	3,647
3	climateaction	3,604
4	climatecrisis	2,399
5	climatestrike	1,634
6	climateemergency	1,289
7	climatefact	1,207
8	climatejustice	574
9	climatebreakdown	423
10	climatefacts	378

Table 3: Most frequent hashtags starting with *clim ranked by frequency**

#climatechange is the most tweeted one, followed by *#climatehope*, and *#climateaction*. However, the most frequent hashtags which describe how climate change is perceived are *#climatechange*, *#climatecrisis* and *#climateemergency*. There is evidence of

how hashtags evolve through time while the climate issues keep worsening, as shown in Figure 2. Indeed, from 2015 to 2017 only #climatechange is recorded and there is a decreasing trend after 2018. Conversely, #climatecrisis and #climateemergency first appeared in 2018. These three hashtags may represent three different levels of an increasing danger perception, where #climatechange seems to be a more neutral and overused expression as to refer to *a fait accompli*. Indeed, the hashtag #climateemergency first appeared on 12th August 2018 posted by the environmental movement Extinction Rebellion, a global non-violent direct-action movement demanding a response to the climate and ecological emergency. #climatecrisis is widely used in the U.S. with a peak in 2019. However, these expressions have become familiar and widely shared.

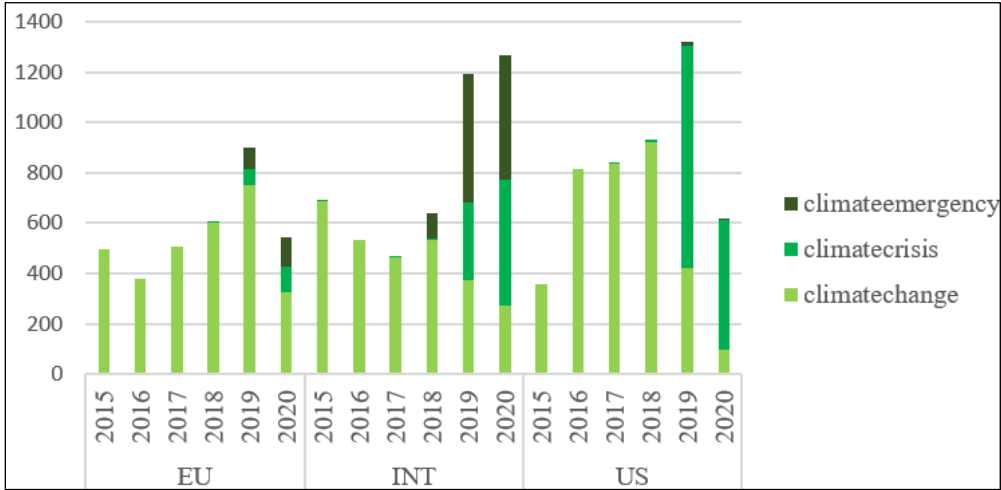


Figure 2: #climatechange, #climatecrisis and #climateemergency per frequency over place and time

In order to determine whether lexical creativity is more frequent in certain groups and if it has a negative or positive meaning through time, hashtags with a range frequency between thirty-two and four, created with two morphemes and starting with *clim**, have been selected manually. These expressions were checked on the *Merriam-Webster Dictionary* to verify whether they have not been accepted as new entries, either as a separate set of words, as in the case of *climate denier*, or as a single-stress word.

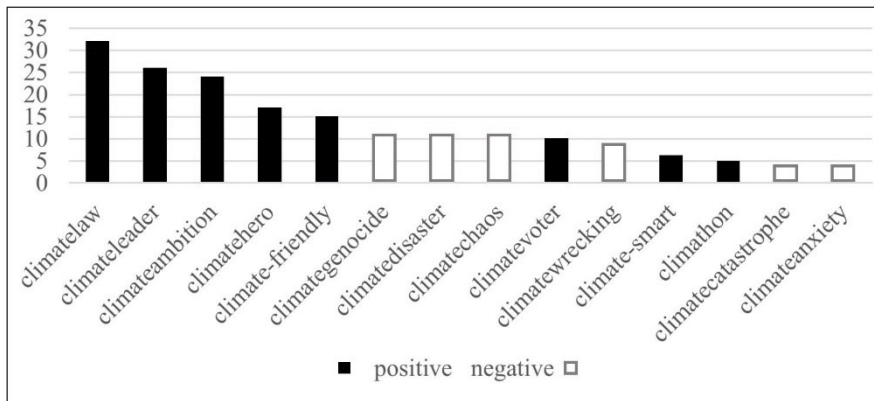


Figure 3: Least frequent hashtags per frequency. The colors refer to the positive or negative connotation of the hashtag carried out manually

Some examples are provided for an in-depth analysis of each new lexical item, its frequency in the corpus and how it is shared among the users. The sample tweets offer a better understanding of the context in which the hashtags occur.

#climatelaw appears thirty-two times, and seventy-three times as #EUClimateLaw. Its first appearance dates to June 2018 used by GreensEFA, a European organization, then also used by politicians such as Mairead McGuinness, Kadri Simson, Frans Timmermans, the activist Greta Thunberg, and the European and international NGOs Green_Europe and Greenpeace.

- (1) *EPIC NEWS! The Netherlands will have the most ambitions #climatelaw in the world!! ??????? we're so glad and proud that our friends at @groenlinks made it happen!!* (GreensEFA 06.2018)

#climateleader is twenty-six times frequent. It was first tweeted in April 2016 by Earth Guardians, an American social movement, then also used by US and EU social movements, NGO, Activist, U.S. organization, i.e. Climate Reality, Extinction Rebellion, 350.org Europe, Climate Action Network, Greta Thunberg, NOAA, The Nature Conservancy.

- (2) *Climate leaders don't just talk. They act. Join us!! Global climate strike 14 December. Spread the word!! #FridaysForFuture #ClimateStrike #ClimateChallenge #ClimateLeader #cop24 #schoolstrike4climate.* (Greta Thunberg 12.2018)

#climateambition appears twenty-four times, the first time in the corpus in May 2015 on Climate Reality's account. Other author types are U.S. and EU organizations, activist, US and EU social movements, NGO, politician, and the users are EU Climate Action, The Nature Conservancy, Greta Thunberg, 350.org Europe, Climate Action Network, Charles Michel.

- (3) *Let's make a collective resolution to ramp up our #climateambition this year.* (The Nature Conservancy 01.2020)

It led to #climateambitions summit in 2020.

#(my)climatehero(s) is seventeen times frequent. The first appearance was in April 2015 with Greenpeace. Thereafter it was used by the European organization EU Climate Action, the American social movements Earth Guardians and Extinction Rebellion, and the NGO Climate Action Network

- (4) *OPEN CALL for youth art! Post your depiction of a hero addressing climate change w/ #MyClimateHero.* (Earth Guardians 04.2017)

#climatefriendly appears fifteen times and the first time was in March 2015 with EcoWatch. Successively, it was also used by Greenpeace, the organization EU Climate Action, and by the politicians Miguel Arias Cañete, Andreea Strachinescu.

- (5) *Getting ready for the #NewYear? Here are some tips to make it more #climatefriendly.* (EUClimateAction 12.2017)

#climategenocide is eleven times frequent. It was used for the first time in September 2018 by the American social movement Extinction Rebellion.

- (6) *“Animals are now dying at a faster rate than during the previous #MassExtinction”; things are only going to get worse as our climate wrecking government supports #HeathrowExpansion; willful #ClimateGenocide??* (Extinction Rebellion 09.2018)

#climatedisaster also appears eleven times starting from April 2018 with Extinction Rebellion, then used by Greenpeace and EcoWatch, as well.

- (7) *The Most Affected People and Areas take the worst of the #ClimateDisaster. They aren't centred in the solutions.* (Extinction Rebellion 11.2020)

#climatechaos was used eleven times. In January 2018 it was first posted by Earth Guardians in the US and then by Extinction Rebellion and 350.org Europe.

- (8) *The kids are the ones who are going to suffer from #ClimateChaos – that's why @solarcentury are backing the youth strikers by supporting the Global #ClimateStrike?? Pledge to join.* (350Europe 09.2019)

#climatevoter(s) was detected ten times. It first appeared in October 2018 with EcoWatch. It was then also used by American activist Jerome Foster II.

- (9) *We have the power to influence this election. We have the power to vote in massive numbers. We have the power to unite as young people. We have the power be #ClimateVoters. Are you #OneOfTheMillion?* (Jerome Foster II 08.2020)

#climatewrecking was posted nine times by American social movement Extinction Rebellion starting from August 2018.

- (10) *Chomsky on the Anthropocene, it causes and solutions. Come to the pre-rebellion warm-up? road block against the #ClimateWrecking UK government today at 7 pm in Westminster. (Extinction Rebellion 09.2018)*

#climatesmart was detected six times starting from April 2016. It was first used by American social movement Climate Reality, and successively by the organization EU Climate Action.

- (11) *#Climateaction makes us #happy?? Check out our tips for #climate-smart everyday action. (EUClimateAction 03.2017)*

#climathon was used five times only by American social movement Climate Reality from June 2015.

- (12) *Drive #ClimateAction in your city! Follow @GlobalClimathon on Twitter and tell your friends about the global #Climathon movement. (Climate Reality 10.2017)*

#climatecatastrophe was tweeted four times by American social movement Extinction Rebellion from October 2018.

- (13) *“Stopping traffic, disrupting people... it is 4 a reason. People need to know the severity of the #ClimateCatastrophe we are facing... if we come together... we can face this crisis together.” Declare a #rebellion w/ us, 31 Oct? (Extinction Rebellion 10.2018)*

#climateanxiety was posted four times, it first appeared in September 2019 with American activist Vic J. Barrett, and successively used by American social movement Extinction Rebellion, and the NGO Greenpeace.

- (14) *Sometimes you just have to be real about what’s going on in your head. Yes I have anxiety, yes I have depression, yes climate change makes it 100x worse. #climateanxiety (Vic J. Barrett 09.2019)*

Overall, these hashtags show whether a sense of connectedness arose among the different users in sharing the same hashtag, also by other author types, and thus by other communities within the corpus, or if they were destined as standalone hashtags within a certain period of time. For instance, hashtags such as *#climateleader*, *#climateambition*, and *#climatehero*, were used by European, American and international users. This proves a common will to act in a positive way towards a shared value. Conversely, *#climategenocide*, *#climaterwrecking*, *#climathon*, and *#climatecatastrophe* were only used by American users. In the other cases, hashtags were used at least by two users from a different origin. Specifically, in the European scenario, in line with the three hashtags analyzed in Figure 2, it appears that the climate change issue is explicitly and formally acknowledged also from politicians. It is not surprising that international users address the topic as an emergency, thus

using more negative hashtags. As for the United States, the data suggest a more complex and discouraging situation where the creation and choice of hashtags may depend on the political context.

Figure 4 refers to the evolution of hashtags through time. *#climateambition* first appeared in 2015 and then again in 2019 and 2020. *#climatelaw* is the most used hashtag from 2016 to 2020. *#climateleader* was also repeated over the years, as well as *#climatehero*, *#climathon* and *#climatechaos*. The usage of the other hashtags is less constant throughout time.

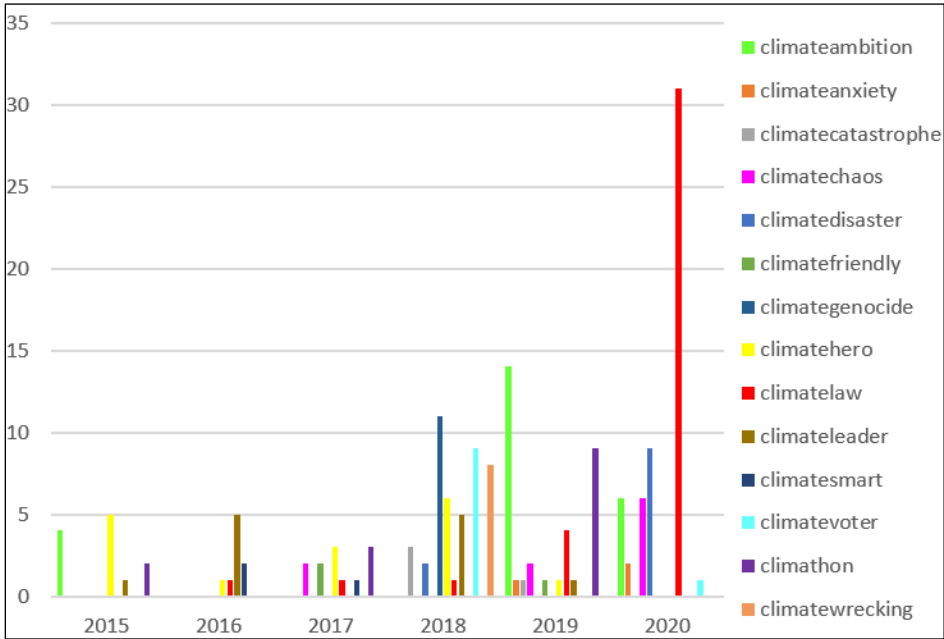


Figure 4: Least frequent hashtags per year

The fast and easy access to this social media platform encourages users to develop a creative language enabling them to make words more suitable to express specific circumstances and feelings, thus fulfilling the function of keeping the virtual community informed. In this case study, from a morphological perspective, the main technique used to form creative hashtags is compounding. As highlighted by Mühlhäusler (2003), the main characteristic of a compound is being created with free-standing lexical forms and the meaning of a compound might differ from the sum of the two lexical forms, such as:

- #climatelaw #climateleader #climateambition #climatehero
- #climategenocide #climatedisaster #climatechaos #climatevoter
- #climatewrecking #climatecatastrophe #climateanxiety #climathon
- #climate-friendly #climate-smart

These hashtags refer to actions, people, feelings, behavior and to an alarmist description of the climate change issue. Indeed, *#climatelaw*, *#climateambition* and *#climathon* refer to actions; *#climateleader*, *#climatehero*, *#climatevoter* to people; *#climateanxiety* represents a new feeling toward climate change perception, *#climate-friendly* and *#climate-smart* are meant as positive behaviors; *#climategenocide*, *#climatedisaster*, *#climatechaos*, *#climatewrecking* and *#climatecatastrophe* are different ways of depicting how the issue is perceived.

As shown in Figure 3, hashtags were manually colored in blue if they conveyed a positive connotation, while in red for a negative one. In a contextualized analysis of these hashtags, the repertoires identified by Ereaut and Segnit (2006: 12) can be employed to better explain the underlying meaning of the tweets. Indeed, in *Warm Words*, Ereaut and Segnit, focusing on a research study on the discourse of climate change in the UK, defined two dominant repertoires in the UK meant as “systems of language that are routinely used for describing and evaluating actions, events and people” highlighting their importance in making sense of an issue. Specifically, they identified a pessimistic category, namely “alarmist”, and a positive one, “small actions”, that is “a more pragmatic and optimistic repertoire”. Indeed, hashtags from the corpus, such as *#climatedisaster*, *#climatecatastrophe* and *#climateanxiety* fit the alarmist repertoire which recurs to the use of language that projects apocalyptic and shocking scenarios. On the other hand, hashtags such as *#climatevoter*, *#climatehero*, *#climate-smart*, lead to call to action referring to what each individual can do to tackle climate change.

With regard to user types, the main hashtag users are American social movements, followed by international NEWS accounts and then by European organizations. As reported in Figure 5, other active users are American organizations and international NGOs. A small contribution is also given by European politicians and European social movements. In particular, American social movements adopted the selected hashtags twelve times more compared to the European ones. There is a slight difference between European and American organizations, forty-one and thirty-one tweets, respectively. Only European politicians used these hashtags. As a result, American users are more active and recur to more alarmist hashtags. Conversely, European hashtags are more positive and are built around more user types.

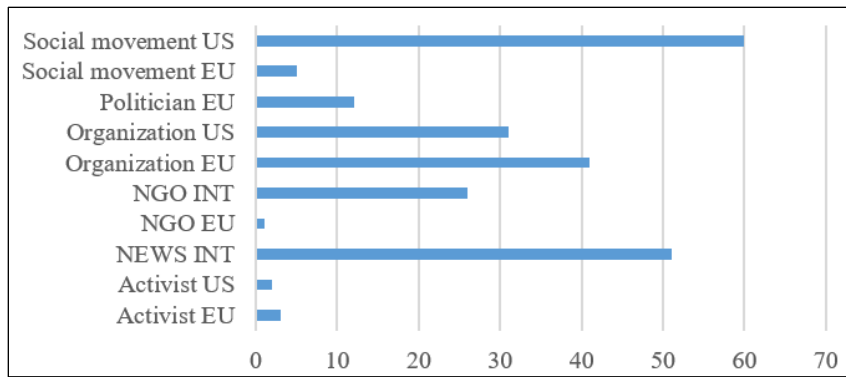


Figure 5: Hashtags per author type (in raw numbers)

5 Conclusions, limitations and future perspectives

Since Twitter, as part of the most used social media platforms, is able to increase the level of connectedness among users, the hashtags addressed above are useful since they express how people and communities communicate their actions, feelings and perceptions about climate change. The possibility to manifest such features through hashtags, strengthens an ideal characteristic of hashtags of acting as a social glue, in the metaphorical sense as highlighted by Klein (2003), that indicates a mutual bond of recognition, connectedness, feeling of responsibility for, and concern for others because they occupy the same community. Indeed, the analyzed hashtags either promote a pessimistic picture of the situation, both defining a state-of-art, and feelings, or call on an optimistic action plan with the aim of involving as many as possible. This effort to raise awareness, take action and involve other individuals is highly shared among the different users from the two geographical areas, and also supported by international users, promoting the sense of connectedness and unity.

As for the first hypothesis, lexical creativity is more frequent in three groups, i.e. American social movements, international news sites and European organizations. There is an explicit call for action providing the opportunity to re-imagine how we live; highlighting how important it is to intervene and what would be the beneficial impacts of a possible intervention. American social movements are the most active users in promoting new compounds in hashtags, mainly with a negative meaning. International news sites follow this trend, with a negative meaning as well. However, European organizations have greatly contributed to the diffusion of certain hashtags with a positive meaning, presumably supported by other users such as European politicians and social movements. European politicians are more likely to support a positive action against climate change which is not

true in the American context. Therefore, it is easily understood that occurrences may vary under the impact of political leaders who most likely dictate the trend. This might be connected to the fact that American social movements are extremely active in defining the issue in different shades trying to raise awareness. As pointed out by Stubbs (1996: 107): “No terms are neutral. Choice of words expresses an ideological position”.

The second hypothesis which presumed that lexical creativity has a stronger negative meaning throughout time can be confirmed. Although most hashtags have a more positive meaning, new hashtags with a negative meaning started rising in 2017 until 2020, while hashtags with a positive meaning are repeated through time. Indeed, in 2015 there are only positive hashtags (*#climateambition*, *#climatehero*, *#climateleader* and *#climathon*). In 2016, again there are only positive hashtags (*#climateleader*, *#climatesmart*, *#climatehero* and *#climatelaw*). In 2017, five hashtags are positive (*#climatefriendly*, *#climatehero*, *#climatelaw*, *#climatesmart* and *#climathon*), however the first hashtag with a negative meaning was tweeted, i.e. *#climatechaos*. In 2018, there are four positive hashtags (*#climatehero*, *#climatelaw*, *#climateleader*, *#climatevoter*) and four negative ones (*#climatecatastrophe*, *#climatedisaster*, *#climategenocide*, *#climatewrecking*). In 2019, six hashtags are positive (*#climateambition*, *#climatehero*, *#climatelaw*, *#climateleader*, *#climatefriendly*, and *#climathon*) and 3 are negative (*#climateanxiety*, *#climatecatastrophe*, *#climatechaos*). In 2020, three hashtags are positive (*#climatelaw*, *#climateambition*, *#climatevoter*) and three are negative (*#climateanxiety*, *#climatechaos*, *#climatedisaster*).

To conclude, this case study was an attempt to show a general trend of how language evolves on Twitter with specific reference to climate change. However, it is limited to the given time frame, 2015-2020, and to the selected stakeholders and hashtags. This analysis will be broadened to the analysis of language creativity in other hashtags, also considering the co-occurrence with other hashtags.

Notes

¹ For more information: <https://change4climate.eu/>

References

- Boyd, D. (2010) ‘Social network sites as networked publics: Affordances, dynamics, and implications.’ In: Papacharissi, Z. (ed.) *A Networked Self: Identity, Community, and Culture on Social Network Sites*. New York: Routledge. 39–58.
- Bruns, A. and Burgess, J. E. (2011) ‘The use of Twitter hashtags in the formation of ad hoc publics.’ In: *Proceedings of the 6th European Consortium for Political Research (ECPR) General*

- Conference 2011*, University of Iceland, Reykjavik. Colchester: European Consortium for Political Research. 1–9.
- Cougnon, L.-A., de Viron, L. and Watrin, P. (2022, submitted) ‘Collection of Twitter corpora for human and social sciences: Sampling methodology and evaluation.’ Paper submitted to the *LREC conference 2022*, 13th Edition of its Language Resources and Evaluation Conference of the ELRA, 21-22-23 June 2022. <https://doi.org/10.31235/osf.io/puhw8>
- Crystal, D. (2006) *Language and the Internet* (2nd ed.). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511487002>
- De Cock, B. and Pizarro Pedraza, A. (2018) ‘From expressing solidarity to mocking on Twitter: Pragmatic functions of hashtags starting with #jesuis across languages.’ *Language in Society* 47(2), 197–217. <https://doi.org/10.1017/S0047404518000052>
- Ereaut, G. and Segnit, N. (2006) *Warm Words: How Are we Telling the Climate Story and Can we Tell it Better?* Institute for Public Policy Research (IPPR): London.
- Fownes, J. R., Yu, C. and Margolin, D. B. (2018) ‘Twitter and climate change.’ *Sociology Compass* 12(6), e12587. <https://doi.org/10.1111/soc4.12587>
- Gerrig, R. and Gibbs, R. J. (1988) ‘Beyond the lexicon: Creativity in language production.’ *Metaphor and Symbolic Activity* 3(1), 1–19. https://doi.org/10.1207/s15327868ms0301_1
- Gupta, N. (1992) *Creativity and Values: Educational Perspectives*. New Delhi: Arya Book Depot.
- Guslyakova, A., Valeeva, N. and Vatkova, O. (2020) ‘“Green neology” and its role in environmental education (a media discourse approach).’ *E3S Web of Conferences*. 169. 05004. <https://doi.org/10.1051/e3sconf/202016905004>
- Hajer, M. and Versteeg, W. (2005) ‘A decade of discourse analysis of environmental politics: Achievements, challenges, perspectives.’ *Journal of Environmental Policy and Planning* 7(3), 175184. <https://doi.org/10.1080/15239080500339646>
- Halliday, M. (2001) ‘New ways of meaning: The challenge to applied linguistics.’ In: Fill, A. and Mühlhäusler, P. (eds) *The Ecolinguistics Reader: Language, Ecology, and Environment*. London: Continuum. 175–202.
- Hulme, M. (2009) *Why we Disagree about Climate Change: Understanding Controversy, Inaction and Opportunity*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511841200>
- Jang, S. M. and Hart, P. S. (2015) ‘Polarized frames on “climate change” and “global warming” across countries and states: Evidence from Twitter big data.’ *Global Environmental Change* 32, 11–17. <https://doi.org/10.1016/j.gloenvcha.2015.02.010>
- Killingsworth, M. J. and Palmer, J. S. (1992) *Ecospeak: Rhetoric and Environmental Politics in America*. Carbondale, IL: Southern Illinois University Press.
- Kirilenko, A. P., Molodtsova, T. and Stepchenkova, S. O. (2015) ‘People as sensors: Mass media and local temperature influence climate change discussion on Twitter.’ *Global Environmental Change* 30, 92–100. <https://doi.org/10.1016/j.gloenvcha.2014.11.003>
- Kirilenko, A. P. and Stepchenkova, S. O. (2014) ‘Public microblogging on climate change: One year of Twitter worldwide.’ *Global Environmental Change* 26, 171–182. <https://doi.org/10.1016/j.gloenvcha.2014.02.008>
- Klein, D. C. (2000) ‘Creating social glue in the community: A psychologist’s view.’ Revision of paper presented at *Rising Tide: Community Development for a Changing World*, 32nd annual conference of the Community Development Society, New Brunswick, Canada, Online document. 9 August 2021. <http://www.humiliationstudies.org/documents/KleinCreatingSocialGlueCommunities.pdf>
- Koteyko, N. and Atanasova, D. (2016) ‘Discourse analysis approaches for assessing climate change communication and media representations.’ In: *Oxford Encyclopedia of Climate Change Communication* Publisher: Oxford University Press Editors: Matthew Nisbet. <https://doi.org/10.1093/acrefore/9780190228620.013.489> <https://www.merriam-webster.com>
- Mühlhäusler, P. (2003) *Language of Environment. Environment of Language. A Course in Ecolinguistics*. London: Battlebridge Publications.

- Nerlich, B. and Koteyko, N. (2009) ‘Compounds, creativity and complexity in climate change communication: the case of ‘carbon indulgences’.’ *Global Environmental Change* 19(3). <https://doi.org/10.1016/j.gloenvcha.2009.03.001>
- Nerlich, B., Koteyko, N. and Brown, B. (2010) ‘Theory and language of climate change communication.’ *WIREs Clim Change* 1, 97–110. <https://doi.org/10.1002/wcc.2>
- Pearce, W., Holmberg, K., Hellsten, I. and Nerlich, B. (2014) ‘Climate change on Twitter: Topics, communities and conversations about the 2013 IPCC working group 1 report.’ *PLoS ONE* 9(4), e94785. <https://doi.org/10.1371/journal.pone.0094785>
- Rey, A. (1988) ‘Dictionnaire et néologie.’ *Actes du colloque Terminologie et technologies nouvelles*, 279–89.
- Roginsky, S. and De Cock, B. (2015) ‘Faire campagne sur Twitter. Modalités d’énonciation et mises en récit des candidats à l’élection européenne.’ *Les Cahiers du Numérique* 11(4), 119–44. <https://doi.org/10.3166/lcn.11.4.119-144>
- Schwell, A. (2015) ‘#MuslimRage: When Western fears meet political humor.’ *Ethnologia Europea* 45(2), 85–101. <https://doi.org/10.16995/ee.1170>
- Stubbs, M. (1996) *Text and Corpus Analysis*. London: Blackwell.
- Veltri, G. A. and Atanasova, D. (2015) ‘Climate change on Twitter: Content, media ecology and information sharing behaviour.’ *Public Understanding of Science* 26(6), 721–737. <https://doi.org/10.1177/0963662515613702>
- Zappavigna, M. (2011) ‘Ambient affiliation: A linguistic perspective on Twitter.’ *New Media & Society* 13(5), 788–806. <https://doi.org/10.1177/1461444810385097>
- Zappavigna, M. (2012) *Discourse of Twitter and Social Media: How we Use Language to Create Affiliation on the Web*. London: Continuum.
- Zappavigna, M. (2014) ‘Ambient affiliation in microblogging: Boarding around the quotidian.’ *Media International Australia* 151, 97–103. <https://doi.org/10.1177/1329878X1415100113>
- Zappavigna, M. (2015) ‘Searchable talk: The linguistic functions of hashtags.’ *Social Semiotics* 25(3), 274–91. <https://doi.org/10.1080/10350330.2014.996948>

Acknowledgement

I thank Louise-Amélie Cougnon and Andrea Pizarro Pedraza for their important contributions to this work.

Funding

I gratefully acknowledge funding from Regione Calabria for the doctoral scholarship.

Vanessa Marcella is Adjunct Professor at the Department of Statistics and Finance at the University of Calabria, Italy. She has a bachelor’s degree in Linguistic Mediation, a master’s degree in Information Technology for Humanities, and is currently pursuing a PhD in Language Analysis and Interdisciplinary studies. Her research focuses on language variation in discourse use in relation to climate change in social-networking-based communications (Twitter).

Address: Vanessa Marcella, Department of Political and Social Sciences, and Department of Culture, Education and Society, Università della Calabria, Via Pietro Bucci, 87036, Rende (CS), Italy. [e-mail: vanessa.marcella@unical.it]

POSSIBILITY MODALS IN ENGLISH TOURISM DISCOURSE: VARIATION ACROSS THREE WEB REGISTERS

Aleksandra Radovanović

<https://doi.org/10.5817/CZ.MUNI.P280-0212-2022-7>

Abstract

The present study explores central modals CAN, COULD, MAY, MIGHT, and the semi-auxiliary verb BE ABLE TO in tourism discourse with a focus on register variation. A small-scale analysis conducted on the tailor-made corpus (c. 131,000 words) of three text types/register (promotional texts, e-newsletters, and managerial responses to guests' reviews) revealed an important variation across registers in the distribution of possibility modals, the degree to which their polysemantic nature is displayed and in their capacity to serve specific metadiscourse functions. The findings clearly indicate that the causes of the observed variation lie in the influence stemming from the differing situational characteristics, especially communicative purposes, interactivity, topic domains, and personal involvement.

Keywords

possibility modals, tourism discourse, register, epistemic possibility, root possibility, metadiscourse

1 Introduction

It has been fifty years since Leech (2004: 72) pointed to a wealth of literature on the modal auxiliary verbs in English, and yet they still are among appealing topics of linguistic enquiry, with the scholarly focus increasingly shifting from the semantic properties of modals to the functions they perform in domain-specific discourses (e.g. Huschová 2015). The present study joins this burgeoning field by addressing tourism discourse, a specialised discourse in which the research on modality is relatively scarce (cf. Radovanović 2020: 277–278). The widespread view reiterated in the studies dealing with the language of tourism in general (Calvi 2010, Maci 2018) or its particular genres (Maci 2007, Manca 2016, Suau-Jiménez 2012, 2019) is that tourism discourse is inherently characterised by persuasion. Tourism industry texts, especially promotional ones, are mainly geared towards moulding and affecting the perceptions, attitudes and actions of the audience (readers) so as to instigate the desired consumer behaviour which will ultimately lead to greater profits for providers of tourism services. In their attempts to maximise the number of prospective tourists turned into actual ones, authors rely on and deliberately choose from an array of linguistic means capable

of bringing about persuasive effects. Modals, especially CAN and WILL, play an important role here (Maci 2007, Manca 2016) attested by their higher frequencies in English tourism discourse than in general language (Maci 2018). In what ways linguistic items aid in achieving the text's communicative purpose can be effectively viewed from the perspective of metadiscourse, which has proven to be an important vehicle for the construction and attainment of persuasion. Seen through the lenses of Hyland's (2005) interpersonal metadiscourse, modals feature among important resources used for intersubjective positioning in tourism discourse (Incelli 2017, Suau-Jiménez 2012, 2019), allowing the writer to convey judgements and overtly align himself/herself with readers (Hyland 2005: 49).

In view of this, the present study focuses on the functional class of modals – possibility modals (cf. Biber et al. 1999: 485) – recognised as belonging to the defining features of “Persuasive/argumentative discourse” (Biber & Kurijan 2007: 116–117). Conceived as a contribution to the study of variation in specialised discourses, the study adopts “the register perspective” (cf. Biber & Conrad 2019) and sets out to explore central modals (CAN, COULD, MAY, MIGHT) and a semantically-related semi-modal (BE ABLE TO) across three web registers of tourism discourse in order to determine whether there are any register-induced variations. Specifically, I look into their distribution, their semantic values and the metadiscoursal functions they perform. However, my aim is not to provide a rigorous account of each and every occurrence, but rather to foreground those facets of modal verb use that may help assess whether it varies on account of the differing register characteristics. To this end, a small-scale analysis was conducted on the corpus comprising three different text types (promotional texts on the destination website, e-newsletters produced by a tourism organisation, and managerial responses to online guests' reviews).

The following section deals with the theoretical and analytic tools applied in the study and then the data and methods used for the analysis are described. The findings are presented and discussed in the next three sections. The final section outlines the main conclusions and points to the limitations of the study.

2 Theoretical underpinnings

This study follows a widely adopted polysemy-based approach to modals (e.g. Palmer 1999, Huddleston & Pullum 2002, Collins 2009) and adopts the binary distinction between epistemic modality and root modality (Coates 1983), which has proven useful in related studies (e.g. Huschová 2014, 2015, Radovanović 2020). With the notions of necessity and

possibility being central to either type, the semantic contrast, also termed intrinsic-extrinsic (Biber et al. 1999), propositional-event (Palmer 2001), ties in with the broad distinction between two domains of human activity: cognition and action. Epistemic modality deals with knowledge, belief, truth, etc. related to the status of the proposition, while root modality relates directly to actions by others and by the speaker himself/herself (Coates 1983, 1995, Palmer 1990, 2001). Hence, the former is concerned with the speaker's or writer's (writer's hereinafter) attitude to the truth-value or factuality of a proposition that describes the event or situation and the latter centres around the potentiality of the event signalled by the proposition (Palmer 2001, 2003). Accordingly, epistemic possibility modals convey the writer's "lack of confidence in the proposition expressed" (Coates 1995: 55) and root possibility modals refer to the factors influencing the actualization of the event. The two types differ in terms of scope. Epistemic possibility takes scope over the whole proposition, while root possibility limits its scope to the verbal element only (cf. Huschová 2014), which is further indicated by appropriate paraphrases ('*it is possible that...*' implies an epistemic reading, while '*it is possible for...*' or *to+inf.* construction reveals a root one). Further, epistemic modals have been attributed certain syntactic restrictions and distributional properties highly uncharacteristic of root modals (taking the perfect/progressive aspect as their complements, taking inanimates, pleonastic/expletive *it* and existential *there* as their subjects, co-occurrence with stative verbs (Coates 1983, Palmer 1990, Collins 2009).

Unlike epistemic possibility, root possibility is semantically highly heterogeneous. Here, I follow a relatively consensual view (e.g. Coates 1983, but cf. Depraetere 2016 for useful discussion) that root possibility falls into three subcategories and, drawing on Palmer's (1990) terms, include: neutral possibility, subject-oriented possibility, and permission. Permission is most clearly observed in the utterances paraphraseable by *be permitted/allowed to*. In the case of neutral possibility the source of possibility involves external circumstances or conditions, while with subject-oriented possibility the possibility originates in the inherent properties of the subject. The reason for adopting the term *subject-oriented possibility* instead of a more conventional one, *ability*, is that the latter is typically ascribed to animates, while subject orientation is also possible with inanimates (Palmer 1990: 85). The distinctions between the types of root possibility are, however, far from being water-tight, especially with ability, as it always implies possibility (Coates 1983, Leech 2004). In fact, discerning modal meaning unequivocally is quite challenging primarily due to "the fuzziness of the root/epistemic boundary in the expression of possibility" (Coates 1995: 65) coupled with the

lack of contextual cues sufficient to resolve modal ambiguities. This may lead to singling out the class of ambiguous (Biber et al. 1999) or indeterminate (Huschová 2015) occurrences, which I conveniently avoided in the analysis. The somewhat simplistic approach followed here might nonetheless be sufficiently effective as it complies with the set aims. Hence, each occurrence was assigned the meaning considered to be the most salient one. In ruling out possible ambiguities, the concepts related to the pragmatic layer of meaning were taken into account where relevant, illocutionary force in particular (Brown & Levinson 1987).

The more pragmatically-grounded strand of the analysis aimed to shed light on the interpersonal functions of possibility modals was grounded in Hyland's (2005) framework of interpersonal metadiscourse, specifically its interactional dimension, which encompasses the following categories modals may fit into: hedges, boosters, attitude markers, and engagement markers. Hedges and boosters are the epistemic stance categories pertaining to the opposite poles of the epistemic cline (doubt and certainty, respectively) (Holmes 1982). Hedges indicate a lack of the writer's commitment to a proposition and allow him/her to present information as an opinion rather than an accredited fact (Hyland 2005: 52), while boosters allow writers to express the highest level of their commitment. Devices signalling epistemic evaluation also serve to modify the illocutionary force of speech acts (Holmes 1982), producing, in turn, some important discourse effects, notably politeness and indirectness in case of hedging (Brown & Levinson 1987, Fraser 2010). Attitude markers are related to affective/emotional aspects of the authorial stance (Hyland 2005: 53). Engagement markers, on the other hand, concern reader's participation and include devices that explicitly address readers.

3 Data and methods

For the purpose of the present study, I compiled a specialised corpus (c. 131,000 words) which contains three types of tourism-related texts freely available online. The intention was to include the text types differing in relevant "situational characteristics" (Biber & Conrad 2009, Biber & Egbert 2018) while ensuring the comparability of sub-corpora in size and the writing time period. The sub-corpus VLdn (43,998 words) consists of promotional texts (on things to do, places to visit, and things to see) retrieved from the website of the official tourism organisation for London in May 2019. To create MRsp (43,294 words), I downloaded 400 responses from London four-star hotels to guests' reviews posted on Tripadvisor (80 per a rating category based on a 5-point bubble rating scale). I selected not more than five responses to the reviews of different ratings given by one specific hotel

representative so that individual variation would not distort the findings. Managerial responses cover the same three-year time span (January 2018 – December 2020) as do 49 issues of e-newsletter released by the European Tourism Association, which constitute Nwslt (44,034 words). This sub-corpus covers professional-professional interaction, whereas the two others involve professional-customer communication. Besides sharing the written mode, a notable similarity for the three text types is that, like most tourism texts (cf. Calvi 2010), they integrate informational and promotional purposes with a persuasive goal, yet in different ways and to differing extents. They differ from each other with respect to other accompanying purposes they are intended to achieve as well as to other characteristics (participants, topic domain, interactivity, personal involvement) and thus may be rightfully regarded as representative of three different register categories. Following Biber and Egbert's (2018) taxonomy, the sub-corpora may well be assigned, respectively, to the following web (sub-) register categories: informational persuasion, interactive explanation, and informational description.

The study adopted a composite approach that integrates complementary perspectives and methods (quantitative and qualitative, software-assisted and manual) commonly applied in studies of modality markers. Software *AntConc* v 3.5.2 (Anthony 2018) was used to generate concordance lines of each central modal and the tensed forms of a semi-modal in three sub-corpora. After eliminating irrelevant occurrences (MAY denoting a month and the syntactically-motivated uses of BE ABLE TO), the raw frequencies (RF) and those normalised (NF) to numbers of occurrence per thousand words (ptw) were calculated. The retrieved data were then subjected to a two-round manual analysis. Applying the frameworks described above, each identified occurrence was first coded for the (sub)type of modal meaning expressed and then for the metadiscourse function it serves. Finally, the results were combined and analysed quantitatively and qualitatively.

4 Distribution of possibility modals

Quantitative findings indicate that possibility modals are quite common in tourism discourse, averaging out to nearly five (4.88) occurrences per thousand words (ptw). However, modal frequencies presented in Table 1 point to rather different distributions across the three web registers. The frequency of modals in managerial responses is almost twice higher than that in promotional texts (3.34 vs 6.14 ptw), while in newsletters it is somewhere between the two (4.77 ptw).

Modal	VLdn		MRsp		Nwslt	
	RF	NF	RF	NF	RF	NF
CAN	126	2.86	159	3.67	150	3.41
COULD	7	0.16	39	0.90	12	0.27
BE ABLE TO	4	0.09	33	0.79	5	0.11
MAY	10	0.23	29	0.67	37	0.84
MIGHT	2	0.05	6	0.14	6	0.17
Total	149	3.34	266	6.14	210	4.77

Table 1: Frequency of possibility modals across registers (per thousand words)

Quite unsurprisingly (cf. Biber et al. 1999: 486–490), the findings show that the frequencies of CAN far outstrip those of other possibility modals in each register while MIGHT is generally very rare. Also, the present tense modals (CAN, MAY) are more frequent than their formal past tense counterparts (COULD, MIGHT) in each register. Except for MAY, which tends to be most frequent in Nwslt, each of the modals displays the highest frequency in MRsp. This is especially the case with COULD and BE ABLE TO, which are considerably more common in MRsp than in other registers. Contrary to my expectations, the findings show that each modal is least favoured in VLdn.

5 Meanings of possibility modals

Overall, my findings accord with those of previous analyses in this field (e.g. Biber et al. 1999, Collins 2009, Huschová 2015). They indicate that in tourism discourse, the use of CAN, COULD and BE ABLE TO tends to be confined to the domain of root modality, while MAY and MIGHT occur as markers of both types of modal meaning analysed in turn below. What emerges from the analysis, however, is that the extent to which the investigated items realise their semantic potential differs widely across registers.

5.1 Epistemic possibility

As Table 2 shows, epistemic possibility is quite underrepresented in my data (overall NF 0.42 ptw), and almost exclusively restricted to MAY/MIGHT. While my corpus offers no examples of epistemic CAN, the use of which is restricted to non-affirmative contexts, there is one instance of COULD signalling the writer’s “assumptions or assessments of possibilities” (Coates 1995: 55) in MRsp.

	VLdn	MRsp	Nwslt
Modal	RF	RF	RF
COULD	0	1	0
MAY	6	18	21
MIGHT	1	4	4
Total	7 (4.70%)	23 (8.65%)	25 (11.90%)

Table 2: Epistemic possibility modals across registers

The highest proportion of epistemic uses is found in newsletter issues (11.90% of modal occurrences in this data set). The reason for this may be found in the characteristics of this register. Newsletter writers inform the members of the Association about important events and developments, which occasionally involves providing their own perspective, as is the case with Example 1 with MAY conveying speculative judgement that expresses a degree of doubt (Palmer 2001).

- (1) *Current indication is this **may** change from end of June/July with the establishment of “travel corridors” possible.* (Nwslt)
- (2) *We are happy to discuss your case further, if we **might** have caused any inconvenience.* (MRsp)

While the examples like Example 2, which refer to the past time, are rather infrequent in Nwslt (two instances of MAY), they constitute almost one third (eight examples) of the modals with epistemic reading found in MRsp. This is hardly surprising since the latter, ranked second in epistemic possibility coverage (8.65%), gives greater prominence to past situations than the other registers. To maintain a good rapport with hotel guests, managers address the issues raised by reviewers, especially in the case of negative reviews, and this is where epistemic modals tend to occur most frequently. On the other hand, in VLdn epistemic possibility seems to be strongly associated with the recurrent notices informing the readers of possible additional financial liabilities (four examples of *additional charges may apply*). In Nwslt, however, instead of MAY, MIGHT tends to be used for this purpose (*a fee might apply*). Considering that we find a 50/50 assessment of possibility with both modals (Coates 1983), this might be seen as a matter of the writer’s personal stylistic preference. The observation that possibility modals are least liable to epistemic interpretation in promotional texts (4.70%) may be explained by the lack of personal involvement coupled with the primarily persuasive communicative purpose of the register. In realising their persuasive goals, the copywriters

focus on the “theoretically conceivable happenings” (Leech 2004: 82) the readers will expectedly be a part of, i.e. the potentialities within the realm of root possibility.

5.2 Root possibility

Table 3 provides the breakdown of root possibility in my data. As can be seen, neutral possibility overwhelmingly dominates over the other subtypes of root meaning in each register. The table also shows that all investigated items readily serve as root possibility markers, but two modals within certain constraints: MAY and MIGHT are incompatible with subject-oriented possibility, and the latter is not used as a permission modal.

Modal meaning		VLdn	MRsp	Nwslt
	Modal	RF	RF	RF
Neutral possibility	CAN	109	124	130
	COULD	6	32	10
	BE ABLE TO	4	20	2
	MAY	4	9	13
	MIGHT	1	2	2
	Total	124 (83.22%)	187 (70.30%)	157 (74.76%)
Subject-oriented possibility	CAN	12	35	15
	COULD	1	3	2
	BE ABLE TO	0	12	2
	Total	13 (8.72%)	50 (18.79%)	19 (9.05%)
Permission	CAN	5	0	5
	COULD	0	3	0
	BE ABLE TO	0	1	1
	MAY	0	2	3
	Total	5 (3.36%)	6 (2.25%)	9 (4.28%)

Table 3: Root possibility across registers

Across investigated registers, expressing neutral, or theoretical (Leech 2004: 82–83), possibility turns out to be the primary use of CAN, COULD and BE ABLE TO and a quite common one of MAY and MIGHT. In this regard, MRsp clearly stands out for the highest frequencies of neutral COULD and neutral BE ABLE TO. This can be partially explained by the above-mentioned bearing of past events on managerial responses, which further entails the use of these items with a past time reference. As Examples 3 and 4 illustrate, COULD indicates that the possibility existed but was not realised, while *were/was able to* implies the actualisation of

the situation. Another reason behind the greatest frequency of BE ABLE TO in MRsp is the tendency of the semi-modal to occur in the syntactic structures far more strongly associated with CAN (e.g. *We hope we will be able to welcome...*), which, however, does not seem to be exclusive of MAY either (e.g. *I sincerely hope that we may welcome you...*).

(3) *We are sorry that, due to Covid-19, you **could** not experience all we have to offer on this occasion.* (MRsp)

(4) *We are glad to read that you **were able to** be back home on time...* (MRsp)

Still, the lowest percentage of neutral modals is found in MRsp (70.30%), while the highest one is noted with VLdn (83.22%). What permeates promotional texts are modal occurrences with the second person pronoun, as in Example 5. Used with the verbs of wanting (*want/wish*), MAY and MIGHT indicate a course of action to the reader, in which respect, they resemble CAN/COULD in their “implicative uses” (Palmer 1990) to make suggestions.

(5) *or depending on your cravings you **might** go for a Trader Vic’s signature curry.* (VLdn)

While *you can* is quite common in Nwslt (e.g. *you can read/access*), sporadic in MRsp, it is pervasive in promotional texts used as an effective means to raise the reader’s awareness of the activities (e.g. *you can watch/travel/tuck into*) and, as Maci points out (2007: 57), expresses “an off-record invitation”. To these ends, the nouns referring to the target groups of readers may be used instead of *you* (e.g. *youngsters/adults/the adventurous can*). VLdn and Nwslt also provide the examples of neutral MAY, such as Example 6, which, according to Coates (1983), might be taken as the instance of “merger”, the coexistence of two meanings (epistemic and neutral) in a both/and relationship. Lastly, this group subsumes the occurrences of CAN, including Example 7, to convey “existential modality” (Palmer 1990), i.e. to indicate “what is sometimes the case” (Huddleston & Pullum 2002: 184).

(6) *This provides a helpful summary of intention to travel and some of the measures that **may** boost confidence based on an extensive online survey from spring 2020.* (Nwslt)

(7) *Finding people at the carnival **can** be tricky with road closures and streets full of revellers.* (VLdn)

Compared to neutral, the possibility of subject-oriented type is far less represented. The instances of modals that, like CAN in Example 8, refer to the skills and capacities of the animate subjects and express meanings closely equivalent to *be capable of* or *know how to* (Leech 2004: 74), i.e. ability, are quite rare and, reasonably enough, found mainly in MRsp. More frequent, and found in each register, are the occurrences of subject-oriented CAN

indicating that inanimates have the necessary properties, or power, to make the situation possible (Palmer 1990), as in Example 9.

- (8) *We hope we will have the opportunity to welcome you back, as we are sure that we **can** wow you in future.* (MRsp)
- (9) *Discover how the Cloud **can** help level the playing field and enable SMEs to compete with the travel industry's 'big players'.* (Nwslt)

Along with this use, CAN/COULD commonly express currently actualised ability (Huddleston & Pullum 2002) co-occurring with the verbs of perception (*see*) and cognition (*think, understand*), i.e. so-called ‘private verbs’ (Palmer 1990: 86–88). These are “only tenuously linked with the concept of ‘Ability’” (Coates 1983: 90) as are the occurrences of CAN with performative verbs (*assure, apologise, confirm*), which also fall under this rubric (cf. Kjellmer’s (2003) discussion of “dummy *can*”).

Permission turns out to be the least significant of root possibility meanings (Table 3). The instances of subjective permission are found only in MRsp, involving MAY and COULD in permission-asking examples. Example 10 shows that these may be requests for permission only seemingly since here the reviewer is not really expected to give permission, but it is rather taken for granted. The remaining examples in the corpus involve the source of authority to grant or deny permission that is external to the speech act. It seems noteworthy that in VLdn CAN is the only modal employed for the purposes of reporting ‘rules and regulations’, as in Example 11, while in Nwslt MAY occurs as its more formal substitute.

- (10) ***May** I also mention that the value of the room type and the breakfast exceeded the upgrade value that was paid.* (MRsp)
- (11) *However, members of the public **cannot** attend the State Opening itself.* (VLdn)

6 Metadiscoursal functions of possibility modals

Since “(e)pistemic modality is always a hedge” (Coates 1983: 49), all occurrences of epistemic possibility in my data can be considered hedges. In tourism discourse, pretty much as elsewhere (cf. Hyland 2005), hedges allow writers to express their views with less confidence or tentatively and thus serve as a kind of responsibility disclaimers if things turn out otherwise. The face-saving purposes appear as the prime motive behind hedging in MRsp, most evident in apologetic speech acts. If we take an apology (exemplified by Example 12) as the act which “directly damages S’s positive face” (Brown & Levinson 1987: 68), hedging then is a self-protective strategy used to redress the managers’ and company’s face.

- (12) *I extend to you my personal apologies for any inconvenience and discomfort which **may** have been caused.* (MRsp)

Current conceptions of hedges give valid grounds for acknowledging the hedging quality of root modality markers as well. Specifically, this applies to neutral possibility modals in the examples balancing between root and epistemic readings: the occurrences of neutral CAN and neutral MAY interpretable as existential modality, or seen as the instances of merger, especially in passives, and those of COULD and MIGHT used as hypothetical modals (cf. Huschová 2015). In Nwslt and less frequently in VLdn, these mainly occur in the illocutions which “predicate some future act” (Brown & Levinson 1987: 65) on the part of the reader (e.g. *can be reviewed/found*) and thus threaten the reader’s negative-face want. In such cases, modals perform the interpersonal negative politeness function (Brown & Levinson 1987: 129-130) in that they minimise the imposition on the reader by toning down the force of potentially face-threatening acts. The same holds for the neutral modals co-occurring with the first person subjects that function as illocutionary force hedges in indirect speech acts, i.e. “hedged performatives” (cf. Fraser 2010). Such examples, all drawn from MRsp, comprise the instances of MAY and COULD used with “pure” performatives in requests to attenuate the force of an illocution (cf. Example 10 above). Likewise, COULD found in indirect requests (e.g. *if you could contact me*) is taken as a hedge. Requests are directive speech acts and directives belong to metadiscoursal engagement markers (Hyland 2005).

As mentioned above, directive illocutionary acts containing *you can* seem to be greatly preferred means of encouraging the audience into desired behaviour, especially in VLdn. Utterances like Example 13 resemble requests in that they are meant to make the reader perform some action, but, unlike requests, they do not serve the writer’s needs primarily. Rather, they attend to the needs of both readers and writers as they direct would-be visitors into the actions they all will benefit from. As such, they are not potentially face-threatening, so there is no need to hedge illocutions.

- (13) ***You can** also try the themed “side-tracks”...* (VLdn)

Instead, *you can* functions here as an interpersonal strategy used to involve the audience and, as Suau-Jiménez (2019) observes, engage readers/customers in the co-creation of values. Hence, possibility modals serve as engagement markers when co-occurring with a reader pronoun *you* in directive illocutionary acts other than requests. While this function is typical of neutral CAN, and far less common for COULD, it may also be ascribed to neutral MAY and MIGHT in their pragmatically specialised uses mentioned above (Example 5).

Finally, MRsp affords the possibility of discerning one more metadiscoursal function fulfilled by CAN, namely that of a booster. Since boosters comprise linguistic resources that emphasise the semantic force of a word or the illocutionary force of a speech act (Fraser 2010, Holmes 1982), I believe the occurrences of CAN in functions of an emphasiser/empathiser, as identified by Kjellmer (2003), may be regarded as metadiscoursal boosters. This applies to the occurrences of CAN with *assure* and *confirm* in MRsp. As Example 14 illustrates, the modal adds further emphasis to the event (taking corrective actions), thereby reinforcing the truth of what is asserted.

(14) *I can assure you that the faults were taken care of now and were resolved.* (MRsp)

Marking solidarity with a reviewer, CAN contributes to attaining the text's main communicative purpose, the maintenance of rapport with an unsatisfied customer. Likewise, *can only* occurring in expressives and representatives (with *apologise*, *(re)assure*, *imagine*) is also taken as a booster on the account that it has the same purpose: by intensifying the illocutionary force, *can only* aids in creating a bond of sympathy and understanding.

The capacity of modals for serving particular functions seems to be register-dependent. Table 4 provides an overview of metadiscoursal functions in my corpus. As shown, website copywriters make extensive use of engagement markers, while newsletter writers rely on hedges most heavily.

		VLdn	MRsp	Nwslt
Metadiscoursal function	Modal	RF	RF	RF
Hedge	CAN	12	3	36
	COULD	2	32	12
	MAY	8	21	24
	MIGHT	0	6	5
	Total	22 (14.77%)	62 (23.30%)	77 (36.67%)
Booster	CAN	0	24 (9.02%)	0
Engagement marker	CAN	63	1	23
	COULD	4	0	0
	MAY	1	0	0
	MIGHT	2	0	0
	Total	70 (46.98%)	1 (0.37%)	23 (10.95%)

Table 4: Metadiscoursal functions of possibility modals across registers

7 Conclusion

This study has explored the use of the modal verbs of possibility and a semantically related semi-modal in three types of tourism texts (promotional texts, managerial responses, e-newsletters) representing three web registers (informational persuasion, interactive explanation, and informational description, respectively). The analysis has revealed certain points of similarity which, however, do not deviate much from the already observed tendencies in terms of occurrences of investigated modals (e.g. the preponderance of CAN, the marginal relevance of MIGHT), their prevailing meanings (e.g. the overwhelming dominance of neutral possibility), and a hedging function. Beyond these similarities, the findings point to a substantial variation across the registers in the distribution of possibility modals, the degree to which their polysemantic nature is revealed and their metadiscoursal capacity. The underlying causes of the observed variation are related to the direct influence from the combination of the situational characteristics of each register. The following is a brief recapitulation.

Promotional texts are meant to convey information so as to convince readers to visit the destination, and thus are overtly persuasive. Here, possibility modals are least common and, accordingly, exhibit quite a limited range of modal meanings. Their use reflects the text's primary goal inasmuch as they tend to occur in directive illocutions referring directly to the reader. The metadiscoursal function of engaging the reader is not characteristic of informational persuasion and, to a lesser extent, has also been noted with informational description, the register intermediate in the use of investigated items. Newsletters are meant to provide the members of the Association with updates pertaining to its activities and topical tourism issues. Although they are mainly focused on descriptive or explanatory information, they pursue persuasive goals in that they are aimed at increasing the readers' participation in the announced forthcoming events. Promotional texts present the elements conducive to the destination attractiveness in a seemingly objective way, while newsletter writers tend to embrace the first-person perspective, which is why possibility modals are more frequently employed here to convey the writer's epistemic attitude and/or as hedging devices. Another important contributory factor to the greater frequency in newsletters is the somewhat interactive part of their nature, with the readers' responses being expected to come in the form of taking some specified action (reviewing, visiting the website) prompted through utterances containing CAN.

The highest density in managerial responses indicates that the use of possibility modals closely corresponds to a degree of interactiveness. Further reasons for the highest

occurrence rates are the topic (i.e. hotel stay), which brings personnel's skills into focus and calls for the increased use of ability modals, and the importance of personal stance. Reasonably, this register displays the greatest array of possibility meanings overall and of each modal. What distinguishes managerial responses from the two other registers is that the writer addresses one specific person, the reviewer, while bearing in mind that the response may influence the visiting intentions of other prospective guests who are reading reviews. The major aims of retaining old customers and attracting new ones entail persuasive purposes. Of prime concern, however, are interpersonal goals. Therefore, the need arises to attend to face wants to a far greater extent than in other registers, which lies behind the uses of analysed modals not found in other registers (e.g. in expressive speech acts such as apologies). Further, these register-specific occurrences have opened a window for associating certain uses of CAN with boosters.

The study, I hope, has thus demonstrated how the use of the investigated items is determined by the situational characteristics of a respective register. It, however, suffers from certain limitations concerning corpus size and a lack of another rater who would engage in the process of coding the data and thus minimise the risk of subjectivity. Finally, the insights this paper offers may be fruitful in raising implications for teaching English for tourism purposes, which, unfortunately, have not been tackled here due to practical constraints.

References

- Anthony, L. (2018) AntConc (Version 3.5.2) [Computer Software]. Tokyo: Waseda University. Online software. 5 March 2019. <https://www.laurenceanthony.net/software>
- Biber, D. and Conrad, S. (2019) *Register, Genre, and Style*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108686136>
- Biber, D. and Egbert, J. (2018) *Register Variation Online*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781316388228>
- Biber, D. and Kurijan, J. (2007) 'Towards a taxonomy of web registers and text types: a multidimensional analysis.' In: Hundt, M., Nesselhauf, N. and Biewer, C. (eds) *Corpus Linguistics and the Web*. Amsterdam: Rodopi. 109–131.
- Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. (1999) *Longman Grammar of Spoken and Written English*. London: Longman.
- Brown, P. and Levinson, S. C. (1987) *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511813085>
- Calvi, M. V. (2010) 'Los géneros discursivos en la lengua del turismo: una propuesta de clasificación.' *Ibérica* 19, 9–32.
- Coates, J. (1983) *The Semantics of the Modal Auxiliaries*. London: Croom Helm.
- Coates, J. (1995) 'The expression of root and epistemic possibility in English.' In: Bybee, J. L. and Fleischman, S. (eds) *Modality in Grammar and Discourse*. Amsterdam and Philadelphia: John Benjamins. 55–66. <https://doi.org/10.1075/tsl.32.04coa>
- Collins, P. C. (2009) *Modals and Quasi-modals in English*. Amsterdam and New York: Rodopi. <https://doi.org/10.1163/9789042029095>
- Depraetere, I. (2016) 'Modality.' In: Riemer, N. (ed.) *The Routledge Handbook of Semantics*. London and New York: Routledge. 370–386.

- Fraser, B. (2010) 'Pragmatic competence: The case of hedging.' In: Kaltenböck, G., Mihatsch, W. and Schneider, S. (eds) *New Approaches to Hedging*. Bingley: Emerald. 15–34. https://doi.org/10.1163/9789004253247_003
- Holmes, J. (1982) 'Expressing doubt and certainty in English.' *RELC journal*, 13(2), 9–28. <https://doi.org/10.1177/003368828201300202>
- Huddleston, R. and Pullum, G. K. (2002) *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781316423530>
- Huschová, P. (2014) 'Possibility readings of *can* and *may* and their potential interchangeability.' *Brno Studies in English* 40(1), 89–104. <https://doi.org/10.5817/BSE2014-1-5>
- Huschová, P. (2015) 'Exploring modal verbs conveying possibility in academic discourse.' *Discourse and Interaction* 8(2), 35–47. <https://doi.org/10.5817/DI2015-2-35>
- Hyland, K. (2005) *Metadiscourse: Exploring Interaction in Writing*. London and New York: Continuum.
- Incelli, E. A. (2017) 'A cross-cultural contrastive analysis of interpersonal markers in promotional discourse in travel agency websites.' In: Palumbo, G. (ed.) *Testi, Corpora, Confronti Interlinguistici: Approcci Qualitativi e Quantitativi*. Trieste: EUT Edizioni Università di Trieste. 65–86. Online document. 15 February 2021. https://www.openstarts.units.it/bitstream/10077/18481/4/Incelli_Testi_corpora.pdf
- Kjellmer, G. (2003) 'A modal shock absorber, empathiser/emphasiser and qualifier.' *International Journal of Corpus Linguistics* 8(2), 145–186. <https://doi.org/10.1075/ijcl.8.2.01kje>
- Leech, G. (2004) *Meaning and the English Verb*. London: Longman.
- Maci, S. M. (2007) 'Virtual touring: The web-language of tourism.' *Linguistica e Filologia* 25, 41–65.
- Maci, S. M. (2018) 'An introduction to English tourism discourse.' *Sociolinguistica* 32(1), 25–42. <https://doi.org/10.1515/soci-2018-0004>
- Manca, E. (2016) *Persuasion in Tourism Discourse. Methodologies and Models*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Palmer, F. R. (1990) *Modality and the English Modals*. London: Longman.
- Palmer, F. R. (2001) *Mood and Modality*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139167178>
- Radovanović, A. (2020) 'The modal expression of necessity in English for tourism.' *ESP Today* 8(2), 275–296. <https://doi.org/10.18485/esptoday.2020.8.2.5>
- Suau-Jiménez, F. (2012) 'Páginas web institucionales de promoción turística: el uso metadiscursivo interpersonal en inglés y español.' In: Sanmartín Sáez, J. (ed.) *Discurso Turístico e Internet*. Madrid: Lingüística Iberoamericana Vervuert. 125–154. <https://doi.org/10.31819/9783865278791-005>
- Suau-Jiménez, F. (2019) 'Engagement of readers/customers in the discourse of e-tourism promotional genres.' In: Sancho Guinda, C. (ed.) *Engagement in Professional Genres*. Amsterdam and Philadelphia: John Benjamins. 341–358. <https://doi.org/10.1075/pbns.301.18sua>

Sources

- <https://www.visitlondon.com> Last accessed 29 May 2019.
- <https://www.tripadvisor.com> Last accessed 30 January 2021.
- <https://www.etoa.org/> Last accessed 23 February 2021.

Aleksandra Radovanović is Assistant Professor at the University of Kragujevac, Faculty of Hotel Management and Tourism in Vrnjačka Banja, Serbia. She is the co-author of two English language textbooks for tourism students. Her main fields of interest are the discourse of English for tourism and hospitality, genre analysis, and contrastive linguistics, specifically tense and modality in English and Serbian.

Address: Aleksandra Radovanović, Faculty of Hotel Management and Tourism, Vrnjačka Banja, University of Kragujevac, Vojvođanska bb, Vrnjačka Banja 36 210, Serbia. [e-mail: aleksandra.radovanovic@kg.ac.rs]

GRAMMATICAL COLLOCATIONS IN ENGLISH EXAM TEXTS WRITTEN BY CZECH SECONDARY-SCHOOL STUDENTS

Věra Sládková

<https://doi.org/10.5817/CZ.MUNI.P280-0212-2022-8>

Abstract

This paper focuses on the frequency and accuracy of five types of grammatical collocations G8E-G8I (Benson et al. 1986) in CZEMATELC, an English language learner corpus (8,338 types; 211,503 tokens) consisting of 1,841 English exam texts from the written part of the national school-leaving exams between 2015 and 2019. The findings reveal the prevalence of A1-A2 CEFR level colligations relying on a limited number of verb lemmas, a wide incorrect pattern variation and preference for patterns which are also the most frequent patterns of their Czech equivalents.

Keywords

post-verbal complementation, grammatical collocations, colligations, non-finite clauses, complementation patterns, two-verb sequences, infinitive clauses, *ing*-clauses

1 Introduction

Verbs “have a central role in language processing” (Viberg 2002: 62) and together with their complementation they are believed to “form the communicative core of utterance” (Altenberg 1993: 227). Hence, it is rather surprising that verbs and post-verbal complementation receive very little explicit attention in course books. In fact, their acquisition is largely left to natural language processing, during which L2 knowledge should be induced on the basis of language exposure. Apart from various gap-fills, which may also raise learners’ awareness of post-verbal complementation, course books at lower proficiency levels tend to limit explicit treatment of verbal valency merely to the infinitive-gerund distinction. However, these constructions are often presented as “structural variants... grouped into a single grammar unit filled with differing syntactic specifications and long lists of verbs grouped according to their complement types” (Petrovitz 2001: 172). As learning allegedly arbitrary lists of verbs has been found counterproductive by many teachers and “errors of verbal complementation are among the most frequent and intractable types of grammatical errors produced by ESL learners of all levels” (Burke 2010: 35), it is vital to ascertain which

particular two-verb sequences in a dependency relation require special attention in teaching at primary and secondary schools in order to achieve naturally-sounding complex sentences.

For this reason, five grammatical collocations (i.e. colligations) consisting of two-verb sequences and relevant to ‘Can Do Statements’ in the *English Grammar Profile* (EGP) at B1 CEFR level were selected for the analysis. This choice reflects the requirements for the national school-leaving exam in the Czech Republic and corresponds to grammatical collocations G8E–G8I according to the *BBI Dictionary Classification of Collocations* by Benson et al. (1986). Table 1 shows them also in a more convenient schematic form devised by Hunston and Francis (2000), which will be used throughout this article.

Benson et al. (1986)			Hunston & Francis (2000)
Code	Pattern	Examples	
G8E	verb + <i>to</i> -inf	<i>he decided to come</i>	V <i>to</i> -inf
G8F	verb + inf without <i>to</i>	<i>we must work</i>	V inf
G8G	verb + <i>v</i> -ing	<i>they enjoy watching TV</i>	V -ing
G8H	verb (trans) + obj + <i>to</i> -inf	<i>she asked me to come</i>	V n <i>to</i> -inf
G8I	verb (trans) + obj + inf without <i>to</i>	<i>we let the children go to the park</i>	V n inf

Table 1: Colligations subjected to current analysis

2 Two-verb sequences from different perspectives

Post-verbal complementation can be approached from several different perspectives: a structural approach based on the notion of transitivity, a semantic approach based on meaning-dependent properties of verbs, and a traditional approach regarding word combinations consisting of verbs and particular grammatical constructions as colligations. The dependency relations between the verb and its complements, thus, become word-preference relations requiring item-specific knowledge (Bourke 2010). This last approach is also in line with more recent usage-based theories (Ellis & Wulff 2015) providing evidence that both L1 and L2 language learners are sensitive to the sequential statistics of word dependencies and have substantial statistical knowledge of the units of language and their phraseological patterning (Ellis 2003).

2.1 A structural approach to colligations consisting of two-verb sequences

The analysis of grammatical collocations G8E-G8I (Benson et al. 1986) involves the analysis of complex sentences with non-finite post-predicate complement clauses in subordinate position in the case of patterns *V to-inf*, *V inf*, *V -ing*, *V n to-inf* and *V n inf*, and also the analysis of finite verb phrases containing central modals, *V inf*.

The *V to-inf* pattern is “the most common pattern for *to*-clauses in all registers” (Biber et al. 2002: 329) and the combination *want* + *to*-clause is also extremely common in conversation (ibid.). The subject of the main clause is also the implied subject of the *to*-clause and represents a direct object of the verb (e.g. *need*, *want*).

The *V inf* pattern is very “rare in all registers” (ibid.) and limited only to a few verbs (e.g. *dare*, *help*, *let*) if it represents a verb followed by a bare infinitive clause (e.g. *The police didn't dare touch them.*) even though it also appears in several “fairly idiomatic phrases (e.g. *make believe*, *let go*, *make do*, *hear tell*)” (Hewings 1999). However, the *V inf* pattern is extremely common, especially in conversation, if it represents central modals (e.g. *can*, *would*, *may*) and their complementation.

The *V -ing* pattern, considerably less frequent than *V to-inf* pattern in both BNC and COCA, represents a verb (e.g. *keep*, *finish*, *suggest*) followed by an *ing*-clause, which is also called a gerund or an *ing*-participle, functioning as a direct object of the verb.

The *V n to-inf* pattern is “moderately common in news and academic prose” (Biber et al. 2002: 329) and represents ditransitive or complex transitive verbs complemented by a noun phrase and a *to*-clause, which functions respectively as a direct/indirect object and object complement of the main verb. The implied subject of the *to*-clause can be: 1) the indirect object of the main clause (e.g. *I told grandma to make me some more.*); 2) the direct object of the main clause (e.g. *He believes the results to be unscientific.*); or 3) the subject of the main clause (e.g. *I promised Billy to take him fishing.*).

The *V n inf* pattern, which represents a verb complemented by a direct object and a bare infinitive clause, is much less common than patterns with *to*-clauses. The range of verbs with this type of complementation is also narrower and includes causative verbs (e.g. *get*, *have*, *make*, *help*, *let*) and verbs of perception (e.g. *feel*, *hear*, *see*, *watch*), which, used in this pattern, emphasize that the action expressed by a bare infinitive is perceived as a whole whereas when used in *V n -ing* pattern, the action is perceived only partially and/or as in progress (Carter & McCarthy 2006). The verb *help* can be used in four different patterns, *V inf*, *V to-inf*, *V n inf*, *V n to-inf*, and apart from passive constructions which exclusively take

to-infinitives, bare infinitives significantly prevail in spoken English and when *to help* directly precedes the non-finite complementation.

The choice of complementation if the verb alternates between the *V to-inf* and the *V -ing* patterns is bound to reflect the differences in meaning, but they may not be distinct and noticeable enough to be cognitively approached by EFL learners with a limited amount of language experience. However, new rules stemming from corpus-based approaches might reveal more tangible collocational preferences. For example, the aspectual verb *start* shows a strong preference for state or change of state verbs in the *V to-inf* pattern and for action verbs in the *V -ing* pattern (Kaleta 2017).

Consequently, it is very difficult “to acquire complement knowledge, even for native speakers” (Kitikanan 2011: 279), especially as “grammarians classify verbal complements unteachable because of their complexity” (ibid.). Verbal complementation is usually taught with the help of lists of verbs and/or patterns that need to be memorised, but these lists tend to be long (see e.g. Biber et al. 2002: 332–333), give the impression of arbitrariness, and contain information which may not be relevant for the particular learners.

Moreover, many lists in grammar books and course books differ in terms of comprehensiveness, probably due to attempted simplification. For example, *English File Pre-intermediate Student’s Book* presents the verbs *forget*, *remember* and *try* only in the *V to-inf* pattern (Latham-Koenig et al. 2012: 158) despite other lists generated by Carter and McCarthy (2006), Gethin (1992) and Hewings (1999). More recent course books, such as *Third Edition Maturita Solution Pre-intermediate Student’s Book* (Falla & Davies 2017), do not present the verbs in lists, but require active involvement from students to create their own lists and develop dictionary skills. This is because lists have a strong potential to be inaccurate, as detected by current corpus linguistics research. Hunston and Francis (2000), for example, report “a certain amount of leakage across the two patterns, and some verbs which are normally thought to have the pattern *V to-inf* only do actually sometimes occur with the other pattern, *V -ing*” (ibid.: 98) speculating that the process of language change driven by analogy might be in progress. Similarly, the preference for bare infinitives after the verb *help* has been ascribed to language change under the influence of American English (McEnery & Xiao 2005).

2.1.1 Post-verbal complementation in Czech

Equivalents of the studied verbs in Czech, including modal verbs, are inflected and marked for tense, and many of them can be complemented by the only infinitive that Czech has. It resembles a bare infinitive and is marked by *-t* (*-ti*) or *-ct* (*-ci*) endings depending either on the register or the previous consonant. Infinitive complementation, however, may not be possible with some verbs (e.g. equivalents of *hope* and *expect*) which can be followed only by finite clauses. According to Dušková (1988), if both finite and non-finite complementation is possible, finite complementation is more frequent.

In terms of form, the closest counterpart of a gerund in Czech is a noun with several features similar to those of a verbal noun in English. Derived directly from the verb using a single suffix *-ní*, it is closely related to the action described by the verb and has the capacity to function syntactically as an object of the main verb. It can be pre- and post-modified, but it does not retain verbal characteristics (i.e. the meaning of the verb cannot be complemented further), so if further complementation is required, the verbal noun must be substituted by a finite clause.

2.2 A semantic approach to verbal complementation

Using the results of structuralist analysis as major sources for the pedagogical treatment of verbal complementation, however, was refused by Bourke (2010), who claimed that formal rules help learners only to a limited extent as many learners seem to operate according to the economy principle and/or use intuition when producing a text. Faster and more reliable processing should be achieved if learners relied on the semantic features of verbs to help them predict the appropriate complementation, which is believed to be meaning-driven as “a difference in syntactic form always spells a difference in meaning” (Bolinger 1968: 127).

Researchers working within the cognitive framework widely accept *to*-infinitive to be bound to the “path image-schema” (Langacker 1991: 446) drawing on the view that the *to*-particle is derived from the prepositional meaning ‘toward’. As such, the *to*-infinitive can be mentally represented as a path leading to the completion of an activity. Hence the notion of futurity arises. The concept of gerund encoding “temporal overlap between the main clause event and the complement clause event” (Kaleta 2017: 94) has been more controversial, so it is also seen as a construction carrying a progressive function or as an entity in the process of nominalization (Langacker 1991).

According to Bolinger's (1968) principle, verbs that express something hypothetical, unfulfilled, and/or carry the notion of unrealized possibility concerning the future (e.g. *want, wish, hope, expect, decide, plan*) tend to be complemented by an infinitive, whereas verbs (e.g. *enjoy, finish, practise*) that express something real, vivid and fulfilled that has already happened or is in progress tend to be followed by a gerund. Karttunen (1971), nevertheless, noticed that the verbs *manage* and *fail* clearly contradict this principle. Among other semantic factors which influence the choice of the complement type, Kiparsky and Kiparska (1970) emphasize factivity and claim that factive verbs (e.g. *regret, admit*) typically complemented by gerunds or *that*-clauses, show the reference to an event which took place before the action of the main verb. On the contrary, most infinitive clauses have future reference in relation to the time of the main verb (e.g. *hope, want, wish*), which is coined as non-factive.

Bourke (2010) explains another semantic concept of actuality and fulfilment, implied by gerundial verbs, while verbs requiring infinitives imply potentiality. He also considers tense and aspect to be semantic features affecting the choice of verbal complementation because past-time reference and durative aspect are associated with a gerund whereas volitional verbs (e.g. *want, wish, refuse*), which carry non-durative aspect, require infinitives as they "refer to the onset of an action" (ibid.: 45).

These semantic and/or partially pragmatic criteria (e.g. context, reference, presupposition, felicity conditions of speech acts), observed mainly by Bourke (2010), appear to provide valuable guidance for the selection of appropriate complementation. However, many learners at lower proficiency levels may be unable to group the verbs and generalize their meanings, or they may not fully understand the concepts of factivity, actuality and time-reference, and they may find the number and specifications of these rules overwhelming.

2.3 Collocational patterns consisting of two-verb sequences

Grammatical structures subjected to usage restrictions due to the meaning of specific lexical items can be best approached from the lexico-grammatical perspective because patterns share some aspects of meaning (Hunston & Francis 1998, 2000) and "lexico-grammatical associations correspond to functions" (Conrad 2005: 398). Therefore, instead of seemingly random lists, learners can learn functional patterns and have at their disposal "semi-preconstructed phrases that constitute single choices" (Sinclair 1991: 110).

Verbal complementation is verb specific because the main verb of the sentence determines which type of complement is permitted to follow it. Hence, the frequency of verbs

in target complement constructions in learners' previous language experiences strongly affects cognitive or abstract representations of complementation patterns in the learners' minds and enables them to create templates for future language use (Bybee 2010). These abstract representations of complex constructions, or constructional schemas, can be at various levels of abstraction: high-level constructional schemas appear in the form of general patterns or general rules whereas low-level constructional schemas tend to be more lexically specific (Langacker 2008). The role of frequency in the process of entrenchment and abstraction has not been completely revealed, but repetition and frequent exposure to real language usage in interaction with perceptual salience and ease of processing have a positive effect on the acquisition of infinitive complements, which are considered high type frequency constructions (Keawchaum & Pongpaibroj 2017). The results are less conclusive in the case of gerundial complements, which are considered low type frequency constructions (*ibid.*).

3 Methodology

3.1 Research aim

Based on the premises that “language is highly patterned” (Römer 2009: 141), i.e. made up of word combinations with limited combinatory potential, the analysis of grammatical collocations G8E-G8I (Benson et al. 1986) was approached from the perspective of Pattern Grammar (Hunston & Francis 2000) claiming that “patterns and lexis are mutually dependent, in that each pattern occurs with a restricted set of lexical items, and each lexical item occurs with a restricted set of patterns” (*ibid.*: 3). This is further explained by Hoey (2005), who states that every word is mentally primed either to attract or to avoid particular language features, so words which frequently occur in each other's environment reflect the psycholinguistic connection of lexis in the brain.

Accordingly, the aims of this study were as follows:

1. To determine the frequency and proportion of the five selected grammatical collocations in the exam texts;
2. To identify the accurate and inaccurate uses of these colligations and to categorise infelicitous uses and errors;
3. To ascertain if there is any relationship between the frequency or accuracy of the studied collocational patterns and the CEFR levels according to EGP;

4. To find out if there are any similarities between the studied grammatical collocations containing verbs and the use of verbal complementation of equivalent verbs in the *Czech National Corpus* (CNC) (Křen et al. 2015).

3.2 Research design

The analysis of the studied grammatical collocations was approached in the sense outlined by Halliday (1992), who views grammar and lexis as the notional ends of a lexico-grammatical continuum, and Hunston and Francis (2000), who admit that grammar could be the starting point of the analysis, although lexis is its main focus. As such the analysis concentrated on individual verbs with a potential to appear in the above-mentioned grammatical collocations, and investigated all the patterns in which they appeared.

Frequency analysis, which can also draw attention to language feature avoidance (Ellis & Barkhuizen 2005), was chosen as the main research method. The identification of the colligations was carried out by means of a freely available online corpus analytical tool, AntConc 3.4.4w (Anthony 2014) and manual frequency counts had to be accompanied by detailed qualitative analysis of the context since learner language is full of inconsistencies and inaccuracies in form-meaning mapping.

3.3 Context and participants

The exam texts constituting CZEMATELC were written in English during the spring sessions of the school-leaving exams from 2015 to 2019 by 923 Czech secondary school students attending various types of secondary schools in the Czech Republic. The majority of the students were 19 and had studied English as their L2 for eleven years. The learners might also have been learning German, French, Spanish or Russian as an L3. However, the metadata are unavailable for legal reasons as confidentiality has to be strictly observed in the case of high-stakes exams.

3.4 Learner corpus

The *Czech Maturita Exam Learner Corpus* (CZEMATELC) is a random cluster sample of learner production consisting of 1,841 exam texts based on ten different assignments and representing 0.82 per cent of all assessed texts collected within five years. According to AntConc 3.4.4w (Anthony 2014), the raw corpus contains 211,503 tokens and

8,338 types, but the real number of different words is likely to be considerably lower due to spelling inconsistencies and a large number of proper nouns used.

The clinically elicited data hold some features of both a general and a focused language sample (Ellis & Barkhuizen 2005) because the students were prompted in Czech to achieve a communicative aim in tasks, and their primary focus was on conveying a message. The tasks had some potential to restrict the choice of language features, but they did not force students to use specific target structures.

3.5 Procedure

Firstly, an alphabetical list of all types was created by means of AntConc 3.4.4w (Anthony 2014). The types that were identified as correctly and incorrectly spelt verbs with the potential to appear in G8E-G8I grammatical collocations were viewed in concordance lines and categorised according to the type of complementation. In case of doubt, the accuracy of patterns was checked with the help of the *Advanced Learner's Dictionary*, a reference corpus *Brown Family (C8 tags)*, and a native speaker. Raw frequencies of all studied verbs and their complementation patterns were recorded and compared to language features that appear at different CEFR levels according to EGP. Then, similarities were sought by means of comparison of proportions of individual patterns used with the verbs. In CNC, they were calculated as a percentage of various patterns in which one lemma appears in a randomized sample of 500 concordance lines, a strategy also used by Keawchaum and Pongpairoj (2017).

4 Results and discussion

The analysis revealed that CZEMATELC contains 5,184 correctly created grammatical collocations G8E-G8I (Benson et al. 1986) that appear with 42 different lemmas. Nevertheless, the total number of lemmas used within the studied patterns is 51 because several verbs (e.g. *want, would like, love, start, stop, recommend, help*) are used in two or three correct patterns. Overall, 55 different verbs had the potential to appear in at least one of the studied colligations, but 13 verbs were not used in two-verb sequences and three of them (*remind, instruct, request*) were used only in incorrect patterns. The majority of verbs (43) appear predominantly in patterns, which are also the most frequent patterns of their Czech equivalents and can be seen in Table 2 together with verbs which prefer different patterns. The *V n* pattern ($n = 2,296$), the second most frequent complementation pattern in

CZEMATELC, is strongly preferred by the verbs *like* (n = 497), *love* (n = 427), *need* (n = 246), *make* (n = 168), *help* (n = 128) and *enjoy* (n = 124).

Patterns in CNC	Verbs that prefer the same pattern in CZEMATELC and CNC	Verbs that prefer different patterns in CZEMATELC
V inf	<i>can, will, could, would, must, should, may, might, shall</i>	-
V inf (CNC) V to-inf (CZEMATELC)	<i>want, would like/love, decide, manage, afford, try, begin</i>	<i>start (V -ing)</i> <i>stop (V n)</i>
V n	<i>need, expect, plan, offer, love, remember, hate, bother, enjoy, miss, keep, practise, risk, help, make, recommend, order, advise, motivate</i>	<i>mind V</i>
V clause	<i>hope, suggest, persuade</i>	<i>seem (V to-inf)</i>
V prep	<i>agree, continue</i>	<i>forget (V n), learn (V n), look forward to (V -ing)</i>
V n inf	<i>let</i>	<i>teach (V n)</i>
V n clause	<i>tell, ask</i>	-

Table 2: Verb patterns in CZEMATELC in comparison to CNC (Křen et al. 2015)

The studied colligations, ordered according to their frequencies, are presented in Table 3 together with the CEFR levels at which they tend to appear in learner production. It can be observed that the *V inf* pattern is by far the most frequent and represents more than 65 per cent of all well-formed two-verb sequences while the *V n inf* and the *V n to-inf* patterns can be described as rare. After a detailed comparison against CEFR level descriptors, A1 colligations constitute 66.1 per cent, A2 colligations 28.5 per cent and B1 colligations only 5.1 per cent of the correct patterns. The most frequent incorrect patterns are *V inf* (n = 277); *V past* (n = 144), i.e. incorrect past-tense marking; *V -ing* (n = 77) and *V* (n = 70), both found mainly with modal verbs, *V to-inf* (n = 52) and *V inf-s* (20), i.e. incorrect third-person marking. However, the range of incorrect patterns is wide with a mode of seven incorrect patterns per one verb, but some verbs, such as *can*, *would like/love* and *will* are used respectively in 14, 12 and 11 incorrect patterns.

Pattern	CEFR levels	Raw frequency	Proportion of the pattern	Lemmas	Most frequent errors
V inf G8F	A1–A2	3,394	65.4%	10	V past (n = 95) V (n = 60) V -ing (n = 50)
V to-inf G8E	A1–A2	1,249	24.1%	18	V inf (n = 155) V past (n = 36) V -ing (n = 17)
V -ing G8G	A2–B1	354	6.8%	16	V inf (n = 122) V past (n = 13)
V n inf G8I	B1	134	2.6%	3	V prep n (n = 21)
V n to-inf G8H	B1	53	1.0%	8	V prep n (n = 20) V n for n (n = 23)
Total		5,184	100%	51	

Table 3: Colligations G8E–G8I (Benson et al. 1986) in CZEMATELC

4.1 *V inf* pattern

Apart from the verb *help* (n = 1), all the other instances of the *V inf* pattern contain modal verbs (n = 3,393) and there are many reasons for the high frequency of this pattern with modals. Morphological simplicity and only one correct complementation option make them easy to use (Mitkovska et al. 2012) when other modality resources are limited (Aijmer 2002), especially as half of the studied exam texts carry strong pragmatic functions (e.g. for apologies or requests) or refer to the future. Therefore, the influence of the task, which can be also observed in Figure 1, is undisputable and can explain the overuse of the modal verbs *will*, *can* and *could* as compared to the *Longman Spoken and Written English Corpus* (Biber et al. 2002) which contains native-speakers' data. The underuse of *may*, *might* and *shall* may point to mother-tongue influence and low-input frequency in previous language experience. The exact equivalent of *may/might* is the least frequent modal in CNC and the frequency of *shall* has declined considerably in native-speaker corpora due to distributional fragmentation and paradigmatic atrophy (Leech et al. 2009).

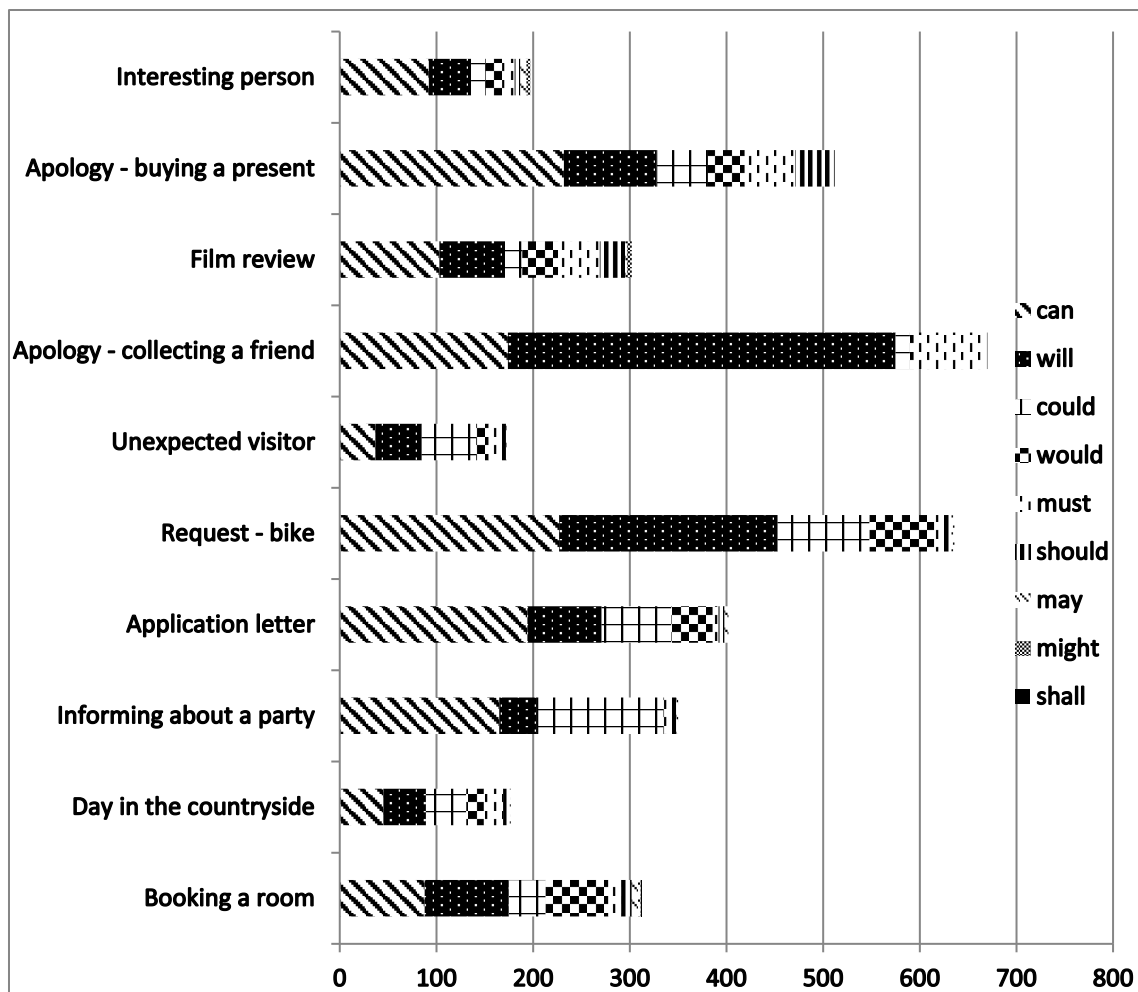


Figure 1: Distribution of modal verbs in individual exam texts

On average, only about nine per cent of modal verbs are used in incorrect patterns and, as seen in Table 4 below, incorrect past-tense marking, i.e. the *V past*, affects the majority of modal verbs and constitutes nearly 30 per cent of all erroneous patterns. It appears predominantly in texts where past-time reference was required (**he was ill and he must learned it*), but also in sentences which clearly refer to the future (**I will not broke it*). Mother-tongue influence can be seen in the *V inf inf* pattern, when modal verbs are complemented by two verbs, usually a lexical verb preceded by an auxiliary verb *be/do* or a modal verb (**On airport will be wait my father*.) and also in the *V* pattern, when complementation is omitted (**We will four people*.). The *V -ing* pattern is exploited mainly by the modal verbs *will*, *can* and *must*, but only patterns with *will* give the impression that the learners had attempted to use future progressive aspect and omitted the auxiliary *be* (**She will waiting to you in park place*). Apart from the *V inf-s* (**she can reads minds*) and the *V to-inf*

(**I could to be satisfy*.), there are twelve other low-frequency patterns, such as *V n -ing* (**You can me writting*) and *V past participle* (**We can seen white deer*).

Verb	CEFR levels	Total	V inf	Errors %	V past	V inf inf	V	V -ing
<i>can</i>	A1	1,357	1,269	6.9	41	7	13	15
<i>will</i>	A1	1,129	990	12.3	19	42	38	21
<i>could</i>	A2	536	498	7.1	22	5	4	1
<i>would</i>	A2	305	287	5.9	2	10	2	1
<i>must</i>	A2	227	205	9.7	5	0	3	10
<i>should</i>	A2	121	110	9.1	5	4	0	1
<i>may</i>	A2	27	24	11.1	1	0	0	1
<i>might</i>	A2	10	9	10.0	0	0	0	0
<i>shall</i>	A2	1	1	0.0	0	0	0	0
Total		3,713	3,393	8.6	95	68	60	50

Table 4: Modal verbs in CZEMATELC

4.2 *V to-inf* pattern

Verbs followed by a *to*-infinitive, which allow other complementation options, are less frequent, and show a higher proportion of incorrect patterns (16.1%) than central modals. For this reason, Table 5 below shows the total frequency of all verbs belonging to a particular lemma, which can be compared to all instances of its clausal complements and correctly created *V to-inf* patterns. The proportion of the pattern in relation to the total number of verbs provides interesting information about learners' choices, especially as it can be compared to the most frequent patterns of their equivalents in Czech based on CNC. The proportion of error patterns in relation to all clausal complements reveals the degree to which clausal complementation of each verb is mastered.

As can be observed, the *V to-inf* pattern was chosen by Czech students frequently when infinitive complementation is also the most or very frequent pattern of the equivalent verb in Czech. When the Czech equivalent does not allow infinitive complementation (e.g. *expect, hope, agree*) or when it is rare (e.g. *forget, learn*), the proportion of *V to-inf* pattern used with this verb in CZEMATELC is also very low. Interestingly, the verb *seem* (n = 8) is predominantly used in the colligation *seem to be* (75%), and the Czech equivalent of this verb, which requires mainly *that*-clauses, allows non-finite complementation only with the verb *be*. The verb *need*, ascribed to A1 CEFR level in EGP, is, in both Czech and CZEMATELC, most

frequently used in the *V n* pattern (n = 246), while the *V to-inf* pattern (n = 46) with this verb belongs to B1 CEFR level. The equivalent of the verb *hope* requires a finite clause in Czech and this complementation is also the most frequent in CZEMATELC (83.9%). The only instances of *hope* in *V to-inf* pattern (n = 15) are conventionalised formulaic sequences which express leave-taking (e.g. *hope to hear from you*). The instances of incorrect pattern *V -ing* (**hope to hearing from you*) could have resulted from contamination with the set phrase *looking forward to hearing from you*.

Verb	CEFR levels	Total	Clausal complements	V to inf		Most frequent pattern in Czech	Errors %
				Raw freq.	%		
<i>would like/love</i>	A1	618	607	559	90.5	V inf	7.9
<i>want</i>	A1	576	467	349	60.5	V inf	27.2
<i>need</i>	A1, B1	339	74	46	13.5	V n	37.8
<i>decide</i>	A2	177	166	124	70.1	V inf	13.3
<i>hope</i>	A2	165	155	15	9.1	V clause	6.5
<i>forget</i>	A2, B1	62	19	6	9.7	V prep	5.3
<i>agree</i>	A2, B2	59	14	4	6.8	V prep	14.3
<i>learn</i>	A1, B1	54	5	3	5.6	V prep	20.0
<i>expect</i>	B1	34	0	0	0.0	V n	0.0
<i>plan</i>	B1	31	13	8	25.8	V n	30.1
<i>offer</i>	A2, B1	31	3	0	0.0	V n	100.0
<i>seem</i>	B1	20	11	8	40.0	V clause	18.2
<i>manage</i>	B1	10	5	5	50.0	V inf	0.0
<i>afford</i>	B1	2	2	2	100.0	V inf	0.0
Total		2,178	1,541	1,129	70.8		16.1

Table 5: Verbs with *to*-infinitive complements

The most frequent incorrect *V inf* pattern (**after school want go England*), which affects mainly the verbs *want*, *would like/love* and *need*, is followed by the *V past* (**I decided to bought*) and *V -ing* (**I want booking rooms*), but altogether there are 18 different erroneous patterns with very low frequencies such as *V to -ing* (**we decided to cycling*), *V to-past* (**we were planning to spent*), *V to clause* (**we are decide to that we meet*), *V past participle* (**I want seen animals*), *V with inf* (**he agree with pick up*), *V sb that -ing* (**I offer my friend that we going*), *V for to-inf* (**you offered for to rent*), *V to n inf* (**she want to him trip up*).

4.3 Verbs alternating between *V to-inf* and *V -ing*

Verbs which can be followed either by *to*-infinitive clause or *ing*-clause take clausal complements in less than one third of instances as many of them prefer complementation by a direct object (e.g. *love, hate, bother*), or a prepositional phrase (e.g. *continue*), which are the prevailing patterns of their Czech equivalents (Table 6). Mother-tongue influence can be observed also in sentences containing the verb *like*, which shows a wide range and a high proportion of incorrectly created patterns (35.1%) although it is expected to be used correctly as early as at A1 CEFR level according to EGP. The most frequent pattern of its Czech equivalent is absent in Table 6 because this verb can be translated in different ways as demonstrated by the following incorrect patterns based on word-for-word translations: *have like + object* (**Terry have got like a history*), *be like* (**I be like from this visitor*), *object + like + subject* (**I hope that this film will like to you*). The problems in word-form mapping concerning this verb might be further supported by confusion of the following entities *look like, would like* (**I'm look like book this accomodation in Hotel Llandough*) and *like* (**I'd like nature in Beskydy*).

Generally, patterns with gerunds prevail over the patterns with *to*-infinitives. In cases when the difference in meaning between the two non-finite clauses is such that only one is mandatory, it is very difficult to identify an incorrect pattern, although there are seven instances of serious doubt. The surrounding co-text tends to be vague, ambiguous and/or contains so many inaccuracies that a clear line cannot be drawn if we do not know what the learner had in mind. For example, we may speculate that the learner meant ‘*we can't stop laughing*’ in the following utterance: *when we are together, we don't stop to laugh*, but the rest of this learner's text is void of any information supporting this speculation.

Verb	CEFR levels	Total	Clausal compl.	V to-inf		V -ing		Most freq. pattern in Czech	Errors %
				Raw freq.	%	Raw freq.	%		
<i>like</i>	A1	688	171	19	2.8	86	12.5	-	35.1
<i>love</i>	A1	509	67	3	0.6	53	10.4	V n	16.4
<i>start</i>	A1	206	122	38	18.4	46	22.3	V inf	31.1
<i>try</i>	A2, B1	84	69	56	66.7	2	2.4	V inf	15.9
<i>stop</i>	A1-B1	45	20	3	4.4	14	28.9	V inf	15.0
<i>remember</i>	A1, A2	45	15	0	0.0	0	0.0	V n	0.0

<i>begin</i>	A1, A2	23	8	1	4.3	1	4.3	V inf	75.0
<i>hate</i>	A2	19	4	0	0	4	21.1	V n	0.0
<i>continue</i>	B1	7	0	0	0.0	0	0.0	V prep	0.0
<i>bother</i>	A2	4	0	0	0.0	0	0.0	V n	0.0
Total		1,630	476	120	7.4	206	12.6		27.1

Table 6: Verbs alternating between V to-inf and V -ing

4.4 *V -ing* pattern

Most verbs followed by gerunds, except for *look forward to*, show relatively low frequencies of *V -ing* pattern in comparison to the total verb frequencies (see Table 7). This pattern seems to be infrequent even in the native-speaker corpus *Brown Family* (C8 tags), in which only one per cent of occurrences of the verb *miss* and 2.5 per cent of the verbs *enjoy*, *suggest* and *keep* appear in *V -ing* patterns. On the contrary, the proportion of this pattern with the verbs *look forward to*, *mind* and *risk* is around 30 per cent. Table 7 shows that in Czech, the equivalents of these verbs appear predominantly in the *V n* pattern, and this is also the case of the verbs *enjoy*, *miss*, *keep* in CZEMATELC. The verb *look forward to* appears most frequently in the *V -ing* pattern, but the *V n* pattern (n = 101) seems to be also relatively common. The high frequency of the well-rehearsed phrase *look forward to hearing from you* (n = 71) might be an entrenched prefabricated sequence, a unique prototypical pattern, acquired by the majority of learners as an invariable chunk. However, learners who achieved a higher level of syntactic complexity were able to use it as a frame and replaced *hearing* with other lexical items (i.e. *seeing*, *meeting*, *receiving*, *reading*). The incorrect patterns (n = 39) might be evidence of U-shaped learning or overgeneralization where the preposition *to* is mistaken for an infinitive particle.

Verb	CEFR levels	Total	Clausal compl.	V -ing		Most freq. pattern in Czech	Errors %
				Raw freq.	%		
<i>look forward to</i>	A2	270	154	115	42.6	V prep	25.3
<i>enjoy</i>	A1	145	20	16	11.0	V n	15.0
<i>miss</i>	A2-B2	48	1	1	2.1	V n	0.0
<i>mind</i>	A2	20	5	5	25.0	V n	20.0
<i>suggest</i>	B1	17	14	4	23.5	V clause	35.7
<i>keep</i>	A2,B1	15	5	5	33.3	V n	0.0

<i>practise</i>	A1	2	1	0	0.0	V n	0.0
<i>risk</i>	B2	2	1	0	0.0	V n	100.0
Total		519	201	146	28.1		24.4

Table 7: Verbs followed by a gerund

4.5 *V n inf* pattern

As seen in Table 8, CZEMATELC contains 134 verbs followed by an object and a bare infinitive, but 115 of these colligations contain the verb *let*. The *V n inf* pattern is the most frequent pattern of its Czech equivalent with a proportion of 65 per cent, and it is also the most frequent pattern of *let* in the *Brown Family* (C8tags). A very low proportion of errors and the fact that 88 instances of this colligation comprise a phrase with a distinct conventional expression (*let me know*), points to an entrenched prefabricated sequence that is only occasionally used as a frame allowing various degrees of substitution ($n = 27$): *let n know*, *let me introduce*, *let me tell you*, *let n inf*. Despite the high frequency of the verb *make*, its colligations in *V n inf* pattern show a low frequency and strong relationship merely with four verbs: *laugh*, *smile*, *feel*, *happen*. The verb *help* is used in three infinitive patterns, *V n to-inf* ($n = 7$), *V n inf* ($n = 4$) and *V inf* ($n = 1$) and also in the ‘*can’t help + ing-clause*’ ($n = 2$). The possibility of gerundial complementation might have influenced some learners to use this verb in an incorrect *V -ing* pattern.

Verb	CEFR levels	Total	Clausal complements	V n inf		Most freq. pattern in Czech	Errors %
				Raw freq.	%		
<i>make</i>	A1–B1	266	21	15	5.6	V n clause	28.6
<i>help</i>	A1, A2	179	16	4	2.2	V n	12.5
<i>let</i>	B1	130	125	115	88.5	V n inf	0.6
Total		575	162	134	23.3		6.2

Table 8: Verbs in V n inf pattern

4.6 *V n to-inf* pattern

CZEMATELC contains 2,331 verbs with the potential to appear in *V n to-inf* pattern, but only 2.3 per cent (i.e. 53 verbs) are really used in this pattern. Many of the 15 lemmas, which can be seen in Table 9 below, namely *would like*, *want*, *help*, *expect*, can take other complements and tend to prefer them. The verbs at higher proficiency levels (e.g. *motivate*,

persuade) show very low frequencies and appear predominantly in *V n* or exceptionally in *V clause* patterns. Although equivalent constructions of many verbs belonging to this group in Czech often require a finite clause after an object, they are used mainly in *V n* pattern in CZEMATELC. The *V n to-inf* pattern is rather infrequent with some verbs even in the native-speaker corpus *Brown Family* (C8 tags). For example, the proportion of this pattern with the verb *recommend* is only one per cent and five per cent with the verbs *want* and *would like* respectively.

CZEMATELC provides several examples of erroneous patterns (n = 9) showing how learners struggle when trying to create complex structures which would enable them to express more complicated ideas. For example, they omit the *to*-particle (**I teach children how live in forest*); substitute it for a preposition (**I want to ask you for lend me your bike*); omit the object (**I want to excuse me*); add an unnecessary object (**I would like you to tell you*), confuse verbs, such as *try* and *persuade* (**I started to try her to keep him*); and limit the *V n to-inf* pattern to *V inf* pattern by omitting verbal morphology (**I ask lend you bicycle*).

Verb	CEFR levels	Total	Clausal compl.	V n to inf		Most freq. patterns in Czech	Errors %
				Raw freq.	%		
<i>would like/love</i>	A1	618	607	3	0.5	V inf	7.9
<i>want</i>	A1	576	467	15	2.6	V inf	27.2
<i>tell</i>	A1, B2	363	171	3	0.8	V (n) clause	4.7
<i>ask</i>	A1-B1	319	181	12	3.8	V (n) clause	2.2
<i>help</i>	A1,A1	179	16	7	3.9	V n	12.5
<i>recommend</i>	B1, B2	163	9	6	3.7	V n	11.1
<i>teach</i>	A1-B2	51	12	6	11.8	V n inf	50.0
<i>expect</i>	B1	34	0	0	0.0	V n	0.0
<i>order</i>	A2	10	0	0	0.0	V n	0.0
<i>advise</i>	B1	7	0	0	0.0	V n	0.0
<i>motivate</i>	C1	5	1	1	20.0	V n	0.0
<i>remind</i>	B1	3	0	0	0.0	V (n) clause	0.0
<i>persuade</i>	B1, B2	1	1	0	0.0	V (n) clause	0.0
<i>instruct</i>	C1, C2	1	0	0	0.0	V n	0.0
<i>request</i>	B1	1	0	0	0.0	V prep	0.0
Total		2,331	1,465	53	2.3		13.4

Table 9: Verbs in *V n to-inf* pattern

5 Conclusion and pedagogical implications

In this study, English exam texts were analysed to describe the frequency and accuracy of five selected grammatical collocations consisting of verbs and their non-finite complements in relation to CEFR levels and post-verbal complementation in Czech. Based on the above presented data, the main findings can be summarized as follows.

A relatively small number of verb lemmas make up all the studied colligations consisting of two-verb sequences in CZEMATELC and only eleven high-frequency verbs account for 90 per cent of these patterns. When ascribed to CEFR levels, the majority of colligations (94.6%) appear to be A1-A2 colligations based predominantly on high-frequency verbs. This could point to L2 learners' restricted lexicon, which is explained by Viberg (2002) on the grounds that verbs require more processing capacity than other lexical words and "even adult L1 speakers favour a small number of very frequent nuclear verbs" (ibid.: 67). The early-acquired verbs owing to high input frequency are prone to be "both overrepresented in terms of frequency and overextended beyond their semantic range" (ibid.: 69) even in advanced L2 learner production.

The most frequent colligation in CZEMATELC, representing nearly two thirds of all patterns, was found to be the *V inf* pattern consisting predominantly of modal verbs and bare infinitives. This might be explained by similarity to Czech, morphosyntactic simplicity of the corresponding pattern, and the higher proportion of texts in which performance of particular speech acts was required. The second most frequent pattern (*V n*) may reveal the prevalence of elementary verb-argument structures, especially as some verbs (e.g. *expect, offer, order*) avoid clausal complementation completely and only appear in this pattern. More complex patterns at B1 CEFR level (*V n inf* and *V n to-inf*) constitute less than three per cent of the studied colligations in the corpus and require explicit pedagogical treatment together with verbs which tend to be used frequently in these patterns by native speakers (e.g. *make, ask, persuade*).

A possible mother-tongue influence can be seen not only in the very high proportion of verbs appearing predominantly in what is the most frequent pattern of their equivalents in the CNC (78.2%), but also in some error patterns, namely the *V* pattern with central modals and the incorrect *V inf* pattern. Incorrect past tense marking and third-person marking have been reported in Spanish (Benali & López 2019) and Korean (Kim & Yoo 2015) learner production, so they could be developmental errors. The highest raw frequency of incorrect patterns was observed with modal verbs, but taken proportionally, verbs alternating between *V to-inf* and *V -ing* patterns show the highest proportion of clausal-complement errors caused

mainly by infelicitous uses of *begin*, *like* and *start*. Some verbs, however, show striking pattern variations, which may suggest that a relatively large number of learners tend to use a random pattern in their language production when two verbs meet. Hence, excessive pattern variation and reliance on simple structures seem to be a more serious problem for Czech learners than the *to*-infinitive versus gerund distinction covered in course books.

Consequently, a change in teaching approaches is desirable. A relatively slow growth of verb lexicon should be accepted, and the pattern accuracy of early-acquired verbs should be emphasised before they are used as “syntactic prototypes” (Viberg 2002: 67) in teaching patterns in which less frequent verbs appear. Instead of processing the input just for meaning, it is vital to make the inconspicuous complementation patterns salient and more easily noticeable by input enhancement methods, such as colour coding and boldfacing in course books (Sharwood Smith 1993), and/or by drawing attention to correct and/or incorrect patterns during various stages of teaching activities, particularly when there is a special emphasis on corrective feedback. To this end, Wajnryb’s (1990) Dictogloss and Tomesello and Harrison’s (1988) Down the Garden-Path Technique could be adapted because they encourage learners to cognitively compare their elicited language production to correct target-language forms. Szudarski and Carter (2016) summarized the research on collocation learning and suggested that “learning more collocates for the same word is more effective than learning single collocates for a large number of different words” (ibid.: 246). They also recommend combining input flood (i.e. artificially increased frequency of a language feature) with input enhancement, and they find contrasting collocations in learners’ L1 and L2 the most effective. This implies that colligations, including non-finite verbal complements, should be taught continuously over longer periods, and that each particular pattern should be integrated within the situational context rather than taught through decontextualized activities. Nevertheless, more research needs to be done concerning both verbal complementation in Czech learner production and the suitability of the suggested techniques for improving non-finite complementation.

References

- Aijmer, K. (2002) ‘Modality in advanced Swedish learners’ written interlanguage.’ In: Granger, S., Huang, J. and Petch-Tyson, S. (eds) *Computer Learner Corpora. Second Language Acquisition and Foreign Language Teaching*. Amsterdam: John Benjamins Publishing. 55–76. <https://doi.org/10.1075/llt.6.07ajj>
- Altenberg, B. (1993) ‘Recurrent verb-complement constructions in the London-Lund corpus.’ In: Aarts, J., de Haan, P. and Oostdijk, N. (eds) *English Language Corpora: Design, Analysis and Exploitation*. Amsterdam: Rodopi. 227–245.
- Anthony, L. (2014) *AntConc*. Tokio: Waseda University. Online. 11 Jan 2022. www.laurenceanthony.net/software/anconc/

- Benali, H. and López, S. (2019) 'Analysing EFL learners' written corpus for a better teaching and learning of gerund and *to*-infinitive structure.' *Journal for Educators, Teachers and Trainers* 10(1), 84–100.
- Benson, M., Benson, E. and Ilson, R. F. (1986) *The BBI Combinatory Dictionary of English*. Amsterdam: John Benjamins. [https://doi.org/10.1075/z.bbi1\(1st\)](https://doi.org/10.1075/z.bbi1(1st))
- Biber, D., Conrad, S. and Leech, G. (2002) *Longman Student Grammar of Spoken and Written English*. Harlow: Pearson.
- Bolinger, D. (1968) 'Entailment and the meaning of structures.' *Glossa* 2, 119–127.
- Bourke, J. M. (2010) 'Verb complementation: A pedagogical challenge.' *Reflection on English Language Teaching* 6(1), 35–50.
- Bybee, J. (2010) *Language, Usage and Cognition*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511750526>
- Cambridge University Press. (2015) *English Grammar Profile*. Online. 8 Jan 2022. <http://englishprofile.org/english-grammar-profile/egp-online>
- Carter, R. and McCarthy, M. (2006) *Cambridge Grammar of English*. Cambridge: Cambridge University Press.
- Conrad, S. (2005) 'Corpus linguistics and teaching.' In: Hinkel, E. (ed.) *Handbook of Research in Second Language Teaching and Learning*. New York: Routledge. 417–434. <https://doi.org/10.4324/9781410612700-32>
- CQPweb v3.2.29. (2008–2016) *Brown Family (C8 tags)*. Online. 4 Jan 2022. <https://cqpweb.lancs.ac.uk/>
- Dušková, L., Knittlová, D., Peprník, J., Strnadová, Z. and Tárníková, J. (1988) *Mluvnice současné angličtiny na pozadí češtiny*. Praha: Academia.
- Ellis, N. C. (2003) 'Construction, chunking and connectionism: The emergence of second language structure.' In: Doughty, C. and Long, M. H. (eds) *Handbook of Second Language Acquisition*. Oxford: Blackwell. 33–68. <https://doi.org/10.1002/9780470756492.ch4>
- Ellis, N. C. and Wulff, S. (2015) 'Usage-based approaches to second language acquisition.' In: VanPatten, B. and Williams, J. (eds) *Theories of Second Language Acquisition: An Introduction*. London: Routledge. 75–93.
- Ellis, R. and Barkhuizen, G. (2005) *Analysing Learner Language*. Oxford: Oxford University Press.
- Falla, T. and Davies, P. A. (2017) *Third Edition: Maturita Solution Pre-intermediate: Student's Book*. Oxford: Oxford University Press.
- Gethin, H. (1992) *Grammar in Context*. Edinburgh: Thomas Nelson.
- Halliday, M. A. K. (1992) 'Language as system and language as instance: The corpus as a theoretical construct.' In: Svartvik, J. (ed.) *Directions in Corpus Linguistics: Proceedings of the Nobel Symposium 82, Stockholm, 4–8 August 199*. Berlin: Mouton de Gruyter. 61–77. <https://doi.org/10.1515/9783110867275.61>
- Hewings, M. (1999) *Advanced Grammar in Use*. Cambridge: Cambridge University Press.
- Hunston, S. and Francis, G. (1998) 'Verbs observed: A corpus-driven pedagogic grammar.' *Applied Linguistics* 19(1), 45–72. <https://doi.org/10.1093/applin/19.1.45>
- Hunston, S. and Francis, G. (2000) *Pattern Grammar: A Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam: John Benjamins. <https://doi.org/10.1075/scl.4>
- Hoey, M. (2005) *Lexical Priming: A New Theory of Words and Language*. Oxon: Routledge.
- Kaleta, A. (2017) 'Corpus-based methods in cognitive semantics: The case of English clausal complementation.' *LAMICUS* 1(1), 88–109.
- Karttunen, L. (1971) *The Logic of English Predicate Complement Constructions*. Austin: Linguistics research center and department of linguistics at the University of Texas.
- Keawchaum, R. and Pongpaioj, N. (2017) 'The role of frequency on the acquisition of L2 English infinitive and gerund complements by L1 Thai learners.' *PASSA* 54, 29–57.
- Kim, J. E. and Yoo, I. W.-H. (2015) 'A corpus-based study of *to*-infinitive errors in Korean college freshmen's writing.' *The Journal of Asia TEFL* 12(4), 37–60. <https://doi.org/10.18823/asiatefl.2015.12.4.2.37>
- Kiparsky, P. and Kiparka, C. (1970) 'Fact.' In: Bierwisch, M. and Heidolph, K. E. (eds) *Progress in Linguistics*. The Hague: Mouton. 143–147. <https://doi.org/10.1515/9783111350219.143>

- Kitikanan, P. (2011) 'Thai EFL students' ability to use the gerunds and *to*-infinitives.' *The International Journal of Humanities* 9(2), 279–290. <https://doi.org/10.18848/1447-9508/CGP/v09i02/43131>
- Křen, M., Cvrček, V., Čapka, T., Čermáková, A., Hnátková, M., Chlumská, L., Jelínek, T., Kovářiková, D., Petkevič, V., Procházka, P., Skoumalová, H., Škrabal, M., Truneček, P., Vondříčka, P. and Zasina, A. (2015) *SYN2015: reprezentativní korpus psané češtiny*. Praha: Ústav Českého národního korpusu. FFUK. Online. 6 October 2021. <http://www.korpus.cz>
- Langacker, R. W. (1991) *Foundations of Cognitive Grammar, Vol. 2: Descriptive Application*. Stanford: Stanford University Press. <https://doi.org/10.1093/acprof:oso/9780195331967.001.0001>
- Langacker, R. W. (2008) *Cognitive Grammar: A Basic Introduction*. Oxford and New York: Oxford University Press.
- Latham-Koenig, C., Oxenden, C. and Seligson, P. (2012) *English File Pre-intermediate: Student's Book*. Oxford: Oxford University Press.
- Leech, G., Hundt, M., Mair, C. and Smith, N. (2009) *Change in Contemporary English*. New York: Cambridge University Press. <https://doi.org/10.1017/CBO9780511642210>
- McEnery, A. and Xiao, Z. (2005) 'Help or help to: What do corpora have to say?' *English Studies* 86(2), 161–187. <https://doi.org/10.1080/0013838042000339880>
- Mitkovska, L., Bužarovska, E. and Kusevska, M. (2012) 'Corpus evidence for the acquisition of modal verbs of obligation by Macedonian learners of English.' In: *7th ELTAM-IATEFL- TESOL Conference: Continuing Professional Development – Opportunities and Challenges*, 27–28 Oct 2012, Skopje, Macedonia.
- Petrovitz, W. (2001) 'The sequencing of verbal-complement structures.' *ELT journal* 55, 172–177. <https://doi.org/10.1093/elt/55.2.172>
- Römer, U. (2009) 'The inseparability of lexis and grammar.' *Annual Review of Cognitive Linguistics* 7, 141–163. <https://doi.org/10.1075/arcl.7.06rom>
- Sharwood Smith, M. (1993) 'Input enhancement in instructed SLA.' *Studies in Second Language Acquisition* 15, 165–179. <https://doi.org/10.1017/S0272263100011943>
- Sinclair, J. (1991) *Corpus, Concordance and Collocation*. Oxford: Oxford University Press.
- Szudarski, P. and Carter, R. (2016) 'The role of input enhancement and input flood in EFL learners' acquisition of collocations.' *International Journal of Applied Linguistics* 26(2), 245–265. <https://doi.org/10.1111/ijal.12092>
- Tomasello, M. and Herron, C. (1988) 'Down the garden path: Inducing and correcting overgeneralization errors in the foreign language classroom.' *Applied Psycholinguistics* 9, 237–246. <https://doi.org/10.1017/S0142716400007827>
- Viberg, Å. (2002) 'Basic verbs in second language acquisition.' *Revue Française de Linguistique Appliquée* 7(2), 61–79. <https://doi.org/10.3917/rfla.072.0061>
- Wajnryb, R. (1990) *Grammar Dictation*. Oxford: Oxford University Press.
- Walter, E. (ed.) (2008) *Advanced Learner's Dictionary*. Cambridge: Cambridge University Press.

Věra Sládková is an English teacher at the Secondary Specialized School and Secondary Vocational School in Kaplice, Czech Republic. Within her postgraduate programme at the Faculty of Education at Masaryk University, she is investigating interlanguage features in English school-leaving exams written by Czech secondary school students. Currently, she focuses on grammatical collocations and pragmalinguistic features.

ORCID ID: 0000-0001-6039-445X

Address: Věra Sládková, Netřebice 35, Velešín 382 32, Czech Republic. [e-mail: 460333@mail.muni.cz]

ASPECTS OF VOCABULARY KNOWLEDGE IN CURRENT LOWER-SECONDARY EFL COURSEBOOKS AND COURSEBOOKS PUBLISHED 20 YEARS AGO

Michaela Trnová

<https://doi.org/10.5817/CZ.MUNI.P280-0212-2022-9>

Abstract

This article presents a content analysis of vocabulary activities in selected lower-secondary coursebooks used in the Czech Republic. The aim was to verify the categorical system, which will be used in further research, and to determine if vocabulary is given more attention in current teaching materials than in those published 20 years ago.

Keywords

vocabulary, coursebooks, vocabulary knowledge, vocabulary activities

1 Introduction

Words are an integral part of the language. Without knowledge of vocabulary and the meaning of individual words and expressions, interpersonal communication would not be possible. “Language which contains grammatical errors is unlikely to be misunderstood in context, but with lexical errors misunderstanding, incomprehension, or in rare cases even offence, are quite likely” (Lewis 1997: 16). Scrivener (1994: 73) states that even beginners can communicate meaning through individual words or expressions, such as *Yesterday. Go disco. And friends. Dancing.*, whereas knowledge of grammar without knowledge of vocabulary is not sufficient. This article describes the basic concepts and methodology of the content analysis of vocabulary activities in EFL lower-secondary coursebooks, aiming to determine whether vocabulary is given more attention in current coursebooks than in coursebooks published 20 years ago. Another goal is to verify the categorical system based on Nation’s (2001) nine aspects of word knowledge. First, vocabulary-related terms are defined, then attention is given to teaching vocabulary and vocabulary activities. Finally, a content analysis of vocabulary activities in selected coursebooks published in 1999-2000 and those in 2019 follows.

2 Basic research concepts

2.1 Vocabulary

Words are building blocks of any language (e.g. Janíková 2005, Haß 2010, Webb & Nation 2017). However, vocabulary, simply defined as words in a language, includes individual words and items that consist of multiple words and that we keep in mind as a whole (Ur 2012). Scrivener (2011) defines vocabulary as individual words (e.g. *dog, green, wash*) or fixed combinations of two or three words (e.g. *stock market*), while lexis is a database of words and all word combinations in our mind. Lexis, therefore, includes individual words, common expressions, i.e. collocations (e.g. *traffic jam*), and longer combinations of words that we commonly use (e.g. *on-the-spot decisions, I'd rather not say*). We call these phrases multi-word units or chunks. Lewis (1993) states that lexis is the basis of language and language consists of grammaticalised lexis, not lexicalised grammar. Lexis includes words, collocations, fixed expressions (e.g. *Good morning*) and semi-fixed expressions (e.g. *Could you pass..., please?*) (Lewis 1997). We distinguish between function words that do not carry meaning (e.g. prepositions, articles, pronouns) and content words (nouns, verbs, adjectives, adverbs). Ur (2012) speaks of grammatical items as closed sets because they are usually immutable, whereas lexical items represent an open set as they are constantly changing, being added or lost. The terms *words, lexical units, lexical items* and *vocabulary* are used interchangeably in this article.

2.2 Vocabulary selection

The choice of vocabulary to be acquired is crucial in foreign language teaching. The selected vocabulary must meet the needs of the learners, and one way to identify such vocabulary is the frequency with which lexical items occur, i.e. how often a word or expression is used orally or in writing, assisted by language corpora (Ur 2012). If we talk about the frequency of occurrence, we start from word families, consisting of headwords, inflexions, and derivations (Webb & Nation 2017). The best-known list of high-frequency words in English is West's *General Service List of English Words* (1953, as quoted in Nation 1990: 22), containing 2,000 headwords. Based on their research of lemmas, which include a headword and its inflexions without derivations, Dang and Webb (2016, as quoted in Webb & Nation 2017: 12–13) created the *Essential Word List (EWL)*. It contains 624 content and 176 function words, which is a goal that beginners should achieve after two years of learning a language. Research shows that if students knew 800 lemmas, it is likely that they would

know 60 per cent of the words they encounter. This could increase to 75 per cent if they knew their inflexions. The content words are given in 50-word sets according to their frequency, so teachers can better integrate them into their teaching. The next step would be to learn the 3,000 most frequent word families (ibid.). Teachers have access to numerous word lists online, such as *Oxford 3000*^{TM1}, a list of 3,000 keywords in English produced by linguists and teachers, and the *English Vocabulary Profile*². In addition, university students can use the *Academic Word List*³.

2.3 Vocabulary in EFL coursebooks

Quality coursebooks provide teachers with many benefits, such as a consistent syllabus and clear structure. Research in various countries shows that coursebooks are essential for teaching English as a foreign language (e.g. Bausch et al. 1997, Römer 2005, Lund 2006, Sikorová 2010). In addition, quality foreign language textbooks offer the right language models and input sources, including vocabulary (Trnová 2017).

In the past, however, vocabulary was not emphasised as it is today. Schmitt et al. (1997) provide the following information on the development of English language coursebooks in terms of vocabulary teaching between 1970 and 1997: English language coursebooks in the 1970s focused on grammar presentation and practice. In the 1980s, more emphasis was placed on vocabulary, for example, in the first editions of *The Cambridge English Course* (1984–1987) and *Headway* (1986–1997), where vocabulary is already part of the course content.

Thus, it is clear that vocabulary has received increased attention in EFL coursebooks in the last 20 years of the 20th century, as confirmed by Morgan and Rinvoluceri (2004) or Hedge (2000), who states that in recent years, interest in vocabulary and teaching vocabulary has grown among publishers and teachers. That is confirmed by coursebooks published at the turn of the millennium, including *Project* (2nd ed.), which states vocabulary and pronunciation in the content. In *Maturita Solutions Pre-Intermediate* (Falla & Davies 2007) and *Maturita Solutions Intermediate* (Falla & Davies 2008), section A is devoted to vocabulary, and at the end of the book, there are *Vocabulary Bank* pages. Vocabulary selection in current English coursebooks is based on thematic areas. While developing vocabulary, the word-formation and relationships between words are used. The expressions that learners learn as a whole are not neglected, either. Usually, one coursebook is designed for 120-140 lessons and provides at least 1,000 new words (Cunningsworth 1995).

2.4 Teaching vocabulary

Vocabulary selection is followed by introducing new vocabulary or increasing vocabulary, as Nation (1990) states. Lewis (1997) lists the following possibilities to present vocabulary: topic, situation, collocations, notion, narration (e.g. for the noun *letter*, we use the collocating verbs *sign, post, receive, reply, check, address*), metaphor (e.g. *He kicked the bucket.*), key words, person (statements beginning with *I*), phonological chunking and grammar. Scrivener (2011) recommends that new vocabulary be put into context, e.g. words related to the same place or event, words that have something in common grammatically and have similar uses, and words that can be used to succeed in a particular task. Finally, Ur (2012: 66) outlines the practical principles in order for words to pass into learners' short-term memory:

- provide the written and spoken form of the word;
- make sure learners understand the meaning of the word;
- introduce the word so that the learners remember it well (e.g. the use of pictures, objects, gestures, facial expressions, or mnemonics).

To practise vocabulary or establish vocabulary, learners repeat lexical items individually and in short conversations, work with them in various activities and games, use them in sentences, short texts, and dialogues (Nation 1990). Lewis (1997) suggests:

- identifying chunks;
- matching parts of collocations, expressions, parts of dialogues;
- completing collocations and expressions;
- categorising lexical items (e.g. formal and informal expressions, positive and negative connotations);
- sequencing lexical items in a logical order;
- deleting (i.e. determining which lexical item does not belong to the other items).

Scrivener (2011) also mentions matching images to lexical items, crossword puzzles, tables, diagrams and memory games. After controlled (restricted) practice, where learners use only selected lexical items, there are activities that give learners some freedom in choosing lexical items in communication. However, there is still an emphasis on accuracy in less controlled (freer) practice (Spratt et al. 2011). Vocabulary is used in longer listening and reading texts, speaking and writing. However, the lexical item becomes part of the long-term memory only if learners encounter it repeatedly. Vocabulary must therefore be regularly consolidated and deepened.

2.5 Vocabulary activities

Lewis (1997) contrasts activities and exercises: Activities with both linguistic and non-linguistic outcomes are usually best done in cooperation in pairs or groups, whereas exercises are usually done individually and have a linguistic focus. However, others may not contrast the two terms. For example, Scrivener (2011: 397) defines an activity as “a single task, exercise or game for students to work on, usually set by the teacher”, Ur (2012) uses tasks and activities, Webb and Nation (2017) speak about vocabulary learning activities. The present text uses the term *vocabulary activities*.

As Janíková (2005) states, vocabulary acquisition in the process of teaching and learning a foreign language is the most demanding and complex area, and the issue of typology or classification of vocabulary activities is one of the most difficult because a clear definition of criteria is complicated, due to many factors that need to be considered. For example, the typology of vocabulary activities according to Janíková (ibid.) is based on functional criteria, i.e. the primary goal (specific subject of practice).

Meara (2014) performed a quantitative analysis of the periodical *Modern Language Journal*, in which he focused on professional articles dealing with vocabulary issues in the years 1916–2010. In a bibliometric analysis, he also addressed the researchers quoted in the articles and found that the main research areas (1981–2000) refer to Paul Nation’s work. Between 2001 and 2010, a new orthodoxy is even emerging, focusing just on Nation’s work. Therefore, the classification of activities in this research is based on Nation (2001, 2013).

2.6 Vocabulary knowledge

We usually talk about vocabulary breadth, i.e. the number of words the learner knows, but the depth of vocabulary knowledge is just as important. Testing the depth of the word will show which areas need more attention, which both teachers and learners can use. It is also helpful to focus on vocabulary activities, examine which are effective, or find out whether they cover all aspects of vocabulary knowledge, which can be beneficial for both teachers, authors of coursebooks, and other teaching materials (Nation & Webb 2011).

According to Milton and Fitzpatrick (2014), the author of the first modern list of components of word knowledge is probably Cronbach (1942, as quoted in Milton & Fitzpatrick 2014: 3), who distinguishes the following components of vocabulary knowledge: generalisation, application, breadth of meaning, the precision of meaning, and availability. Generalisation refers to the definition of a word, application to its adequate use, breadth of meaning refers to the awareness that a word can have several different meanings, and the

precision of meaning refers to the ability to use a word correctly in different situations. Finally, availability refers to the productive use of the word. Richards (1976, as quoted in Milton & Fitzpatrick 2014: 4) proposes eight assumptions about vocabulary knowledge:

1. A native speaker expands his vocabulary further as an adult, while syntax develops little.
2. Knowing a word means knowing the degree of probability of encountering the word in speech or text.
3. Knowing a word means knowing the limitations in using words according to functions and situations.
4. Knowing a word means knowledge of the word within a sentence structure.
5. Knowing a word means knowing the word below the surface and derived words.
6. Knowing a word means knowledge of word associations.
7. Knowing a word means knowing the semantic value of a word.
8. Knowing a word means knowing many different meanings associated with a word.
9. Nation (2001) speaks of three aspects of word knowledge, further divided into three subcategories, which have both receptive and productive aspects (Nation 2013), as shown in Table 1.

Form	Spoken form	R	What does the word sound like?
		P	How is the word pronounced?
	Written form	R	What does the word look like?
		P	How is the word written and spelled?
	Word parts	R	What parts are recognisable in this word?
		P	What word parts are needed to express the meaning?
Meaning	Form and meaning	R	What meaning does this word form signal?
		P	What word form can be used to express this meaning?
	Concept and referents	R	What is included in the concept?
		P	What items can the concept refer to?
	Associations	R	What other words does this make us think of?
		P	What other words could we use instead of this one?
Use	Grammatical functions	R	In what patterns does the word occur?
		P	In what patterns must we use this word?
	Collocations	R	What words or types of words occur with this one?
		P	What words or types of words must we use with this one?
	Constraints on use	R	Where, when, and how often would we expect to meet this word?
		P	Where, when, and how often can we use this word?

Table 1: Nine aspects of word knowledge – receptive and productive aspects (taken from Nation 2013: 49)

For ease of reference, the form, meaning, and use are referred to as categories, whereas the spoken form, written form, word parts, form and meaning, concept and referents, associations, grammatical functions, collocations, and constraints of use are called aspects.

The first category, i.e. the category of form, is represented by the aspects of spoken form, written form, and word parts. The knowledge of spoken form includes the ability to recognise a word when we hear it (receptive knowledge) and the ability to pronounce it clearly, which also includes the correct use of stress if it is a multi-syllable word (productive knowledge) (Nation 2013). The knowledge of written form includes knowledge of spelling. Among other things, Nation (ibid.) recommends connecting spoken forms with written ones at the word, syllable, and phoneme levels, which should develop an understanding that letters represent sounds. The knowledge of the word parts includes the knowledge that words consist of roots, prefixes and suffixes. With increasing language level, knowledge of word parts and word-formation increases, so it is possible to say that word knowledge includes knowledge of words in the word family, e.g. at a lower level we know forms of *mend*, their meanings and uses such as *mends*, *mended* and *mending*. At a higher level, we will already know *mender*, *mendable* and *unmendable* (ibid.). The aspect also includes knowledge of compound nouns.

The category of meaning includes the aspects of form and meaning, concept and referents, and associations. Learners usually think they know a word if they know what it sounds like (spoken form) and what it looks like (written form). However, it is equally necessary for them to associate the forms of a word with its meaning. The strength of this connection determines how quickly the learners recall the meaning when they see or hear the word and how promptly they recall the forms of the word when they wish to express its meaning (ibid.). Therefore, it is vital that learners become familiar with the form and meaning of the word at once and then have the opportunity to encounter the word over time. As for the aspect of concept and referents, words, especially frequent ones, can have different meanings; these can cover more than one page in a dictionary (ibid.). For example, homonyms are words that we write and pronounce in the same way, but they have a different meaning (e.g. *bank*). We can also distinguish between words that we pronounce the same way but write differently (homophones, e.g. *bear* – *bare*) and words that we write the same way but pronounce differently (homographs, e.g. *row*). The last aspect within the category of meaning is represented by associations. The new word or expression is incorporated into the lexicon by establishing relationships to the already acquired vocabulary. These include synonyms (e.g. *beautiful*, *pretty*), antonyms (e.g. *beautiful* vs *ugly*), hyponyms or hyperonyms (*table*, *chair*,

wardrobe vs *furniture*). However, it is not only the relationships between words from a linguistic point of view but also the relationships between words in the human mind.

The third category, i.e. the category of use, includes the aspects of grammatical functions, collocations, and constraints on use. In order to use a word correctly, we must know what kind of word it is and what grammatical schemes the word belongs to. Collocations are words that often occur together (e.g. *to take a holiday*, *heavy rain*), and it is possible to distinguish between strong and weak collocations, such as *watch out* and *watch a video* (Spratt et al. 2011). In this research, this aspect includes collocations as we understand them from the linguistic point of view as well as phrasal verbs and multi-word units (chunks, e.g. *by and large*). Constraints on use result from differences between cultures (e.g. addressing) and styles (e.g. formal, informal). In English, we have to be careful, for example, when using the word *fat*, whereas in other countries, this word is synonymous with *wealth*, and similarly, the word *old* may be considered rude, while in another context, we may consider it synonymous with *wisdom* (Nation 2013).

Finally, it must be noted that Ur (2012) provides eight aspects of vocabulary: (spoken and written) form, meaning: denotation, grammar, collocation, connotation (positive or negative association, e.g. *moist* and *damp*), appropriateness, meaning relationships and word-formation, and Scrivener (2011) offers a list of 32 items stating what knowledge of a lexical item entails. Given that a great deal of vocabulary research in English coursebooks has taken place in Spain, it is appropriate to mention also the twelve vocabulary knowledge dimensions introduced by Jimenéz Catalán (2002, as quoted in Fernández Orío 2014).

3 Research survey methodology

3.1 Research objectives and questions

The research survey aims to map the difference between vocabulary activities in current English language coursebooks for Years 6 and 7 and those published 20 years ago. Another goal is to verify the categorical system, which will be used in subsequent research. Based on the objectives of the study, the following research questions emerged:

1. *Do the selected current English language student's book and workbook for Year 6 offer more vocabulary activities and cover more aspects of vocabulary knowledge than those published in 1999?*
2. *Does the selected current English language student's book for Year 7 offer more vocabulary activities and cover more aspects of vocabulary knowledge than that published in 2000?*

This study follows a research study by Brown (2010), who analysed even units from nine English coursebooks, but no workbooks. As the language level of the coursebooks analysed in this research is A1–A2 according to the *Common European Framework of Reference*⁴, I have not expected all the nine aspects recognized by Nation (2001) to be covered. Furthermore, I have expected that due to the increased interest in vocabulary in recent years (e.g. Hedge 2000, Nation 2013), the current coursebooks (i.e. those published in 2019) will contain more vocabulary activities. Therefore, I have conducted quantitative content analysis, and used a comparative method to compare the current coursebooks and those published in 1999 and 2000.

3.2 Research sample

The basic set consists of EFL coursebooks intended for Years 6 and 7 and includes coursebooks published by international and Czech publishers. For the research, the following books were selected: *Project Explore 1 Student's Book* (Phillips & Shipton 2019), *Project Explore 1 Workbook* (Phillips, Shipton & Trnová 2019), *Project Explore 2 Student's Book* (Wheeldon & Shipton 2019), *Project 2 (2nd ed.) Student's Book* (Hutchinson 1999b), *Project 2 (2nd ed.) Workbook* (Hutchinson 1999a) and *Project 3 Student's Book* (Hutchinson 2000). The selection is not accidental. On the contrary, it is a deliberate selection because *Project* has been widely used in the Czech Republic since the early 1990s, as evidenced by various researchers (e.g. Sikorová 2010, Schmidtová 2017). Another reason is the concept and methodology of *Project* (2nd ed.) and *Project Explore*, which facilitates comparing the coursebooks.

Oxford University Press published the first edition of *Project* by Tom Hutchinson in the 1980s as *Project English*. In the 1990s, it also became popular in the Czech Republic and was even the first foreign English coursebook imported into the Czech Republic. However, since *Project English* has three levels, consisting of eight units and containing grammar and vocabulary above A2 (e.g. modal verbs with the past infinitive), the first edition is not suitable

for comparison. The second edition, published between 1999 and 2002, considers the *Common European Framework of Reference for Languages*. There are five levels (*Project 1–4* and *Plus*), so the coursebook corresponds with the system of education in the Czech Republic more. *Project 1* is intended for Year 5, i.e. the last year of Stage 1, *Project 2–4* and *Project Plus* for Years 6–9, i.e. Stage 2 of elementary school⁵ and the corresponding years of the lower-secondary grammar school. For these reasons, the second edition is more suitable for comparison than the first edition. *Project Explore*⁶ has four levels for Years 6–9 and the corresponding years of the lower-secondary grammar school. In *Project* (2nd ed.), levels 2 and 3 are analysed. In the current *Project Explore*, levels 1 and 2 are analysed.

3.3 Research process

The research survey analysed all the six units in *Project 2* (Hutchinson 1999) and *Project Explore 1* (Phillips et al. 2019), including *Introduction*, sections A-D, songs, sections on pronunciation, pages devoted to *Project* work, English-speaking countries and cross-curricular links. The workbooks were also analysed. Three units from *Project 3* (Hutchinson 2000) and *Project Explore 2* (Wheeldon & Shipton 2019), selected according to the correspondence of the content, especially in terms of topics and vocabulary, were analysed.

3.4 Research unit and categories

The research unit is a *vocabulary activity*, i.e. the activity in which the vocabulary is presented, practised, and used. The activities sought are classified based on Nation's (2001) nine aspects of word knowledge, described in Section 2.6 above. Brown (2010) supplemented the aspects with a brief description of the activities that fulfil them. As part of the pilot research, activities fulfilling individual aspects were specified and supplemented with specific examples of instructions (Trnová 2021). The inclusion of activities in the corresponding aspect of word knowledge results from the focus and primary goal of the activity.

The reliability of the research tool, the categorical system based on Nation's (2001) aspects of word knowledge, was verified through inter-rater reliability. The raters analysed together Units 1–3 in *Project Explore 1*, then worked independently on Units 4–6, with an average agreement of 92 per cent. In Unit 4, the agreement was 90 per cent; in the student's book, the agreement was lower (85.7%) than in the workbook (96%). In Unit 5, the agreement was 92 per cent; in the student's book, the agreement was lower (88.9%) than in the workbook (95.5%). Finally, in Unit 6, the highest agreement was reached (95%), again with a difference between the student's book (91.7%) and the workbook (97.4%) (Trnová 2021).

3.5 Research results

3.5.1 Research question 1

Do the selected current English language student's book and workbook for Year 6 offer more vocabulary activities and cover more aspects of vocabulary knowledge than those published in 1999?

Project Explore 1 and *Project 2* (2nd ed.) are intended for learners in Year 6 or the corresponding year of grammar schools. In *Project 2* (P2), 317 vocabulary activities were identified, while in *Project Explore 1* (PE1), the number of vocabulary activities is higher (379). The number of activities in the student's books is similar, 197 in P2 and 193 in PE1. Nevertheless, the number of activities in P2 is influenced by the *Pronunciation* section after each unit. In addition, the difference lies in the number of vocabulary activities in the workbooks. In P2, the number of vocabulary activities in the workbook (WB) is much lower (120) than in the student's book (SB) (197). In PE1, the number of vocabulary activities in the WB (186) is similar to that in the SB (193). In both SBs, vocabulary is listed in the table of contents. However, in P2 SB, there are no sections called *Vocabulary*, while in PE1 SB, they appear regularly. P2 contains *Introduction* in both SB and WB; PE1 contains *Introduction* only in the SB. On the other hand, PE1 WB contains *Preparation for testing*. It can be stated that PE1 supports mixed-ability classes with activities in the WB marked with one to three asterisks based on difficulty and *Extra* sections in SB.

Coursebook	<i>Project 2</i> (2nd ed.)				<i>Project Explore 1</i>			
	SB	WB	T	%	SB	WB	T	%
Aspect	197	120	317		193	186	379	
Form	79	42	121	38.2	59	44	103	27.2
Spoken form	61	12	73	23.0	44	10	54	14.3
Written form	8	23	31	9.8	6	31	37	9.8
Word parts	10	7	17	5.4	9	3	12	3.2
Meaning	53	45	98	30.9	75	75	150	39.6
Form and meaning	47	41	88	27.8	62	61	121	32.5
Concept and referents	2	0	2	0.6	0	0	0	0
Associations	4	4	8	2.5	13	14	29	7.1
Use	65	33	98	30.9	59	67	126	33.3

Grammatical functions	46	26	72	22.7	41	49	90	23.8
Collocations	19	7	26	8.2	18	18	36	9.5
Constraints on use	0	0	0	0	0	0	0	0

SB – student’s book, WB – workbook, T – number of vocabulary activities in total

Table 2: Comparison of word knowledge in *Project 2* (2nd ed.) and *Project Explore 1*

Table 2 shows the number and percentage of vocabulary activities representing each of the three categories, i.e. form, meaning, and use, and each of the nine aspects of word knowledge according to Nation (2001). The results for the individual aspects will now be presented based on the percentage of vocabulary activities representing the aspects.

The aspect of form and meaning is the most represented aspect in both P2 (27.76%) and PE1 (32.45%). However, there is a significant difference between the coursebooks in the number of activities focused on form and meaning (88 in P2 and 121 in PE1). I believe that the differences stem from the greater emphasis on the multi-sensory approach and from the fact that current coursebooks are developed with greater regard for mixed-ability classes and learners with special educational needs. Lexis is therefore presented in different ways, through different senses. While in P2, vocabulary is introduced in one exercise, which usually offers two to three activities, in PE1, the already acquired vocabulary is activated through the *You first* section (e.g. *Describe a room in your house.*), followed by a picture of the house, with some rooms and pieces of furniture labelled. Other items in the picture need to be labelled, using expressions in the box below (*Complete the labels with the words in the box.*). Vocabulary work continues with activities fulfilling the aspect of spoken form (e.g. *Listen and check.* and/or *Listen and repeat.*) and that of grammatical functions (e.g. *Read the dialogue and complete...*, *Work with a partner. Ask and answer the questions.*). In P2, the vocabulary to be acquired is usually given in a text with some visual support. Activities in which the already acquired vocabulary is activated appear rather rarely, for example, in *Introduction* (e.g. *Look at the pictures. How many things can you name?*). Thus, it is clear that before using vocabulary in a coherent text, learners are more exposed to vocabulary in PE1 than in P2. In both WBs, learners describe black-and-white pictures and, in the case of PE1, also black-and-white photographs (*Label the picture.*), and solve crossword puzzles with pictures or definitions (*Complete the crossword.*). These activities also represent the aspect of written form because learners have to recall the written form of the word. In both SBs and WBs, some

activities fulfil both the aspect of form and meaning and that of grammatical functions. Learners need to recall vocabulary based on visuals, then use it in sentences or texts.

Vocabulary activities focused on the aspect of grammatical functions represent 23.8 per cent in PE1 and 22.7 per cent in P2. According to Nation (2013), many linguists believe that lexis plays an essential role in grammar, and the choice of a verb affects the grammatical structure of the whole sentence. In addition, beginners may not have a good overview of the language system, do not yet understand how the language system works, and therefore it is necessary to involve explicit learning of the patterns in which lexical items feature (Brown 2010). In PE1, activities within the aspect of grammatical functions follow targeted vocabulary work, as described above. In both SBs, learners use vocabulary in short dialogues, speaking and writing. In the WBs, it is a question of completing or using vocabulary in sentences or texts. The difference lies in the number of activities in the WB, which is higher in PE1 WB than in PE1 SB.

The aspect of spoken form is represented less in PE1 (14.3%) than in P2 (23.0%). The *Pronunciation* section after each unit causes the high percentage in P2. It contains three to four exercises that are further divided into activities (e.g. *Listen and repeat. Which syllable has the stress? Listen and check. Where can you hear the sound? Tick the words. Put the words in the correct column.*). In PE1, pronunciation is included in the units and is thus better linked to the subject matter in sections A–D, which is certainly more appropriate. In both coursebooks, there are activities focused on the correct pronunciation of individual phonemes, pairs of phonemes, and the pronunciation of vocabulary intended for acquisition. In P2, there are activities focused on word stress, while in PE1, there are activities focused on sentence stress, but these are not only activities in the *Pronunciation* section. Another difference lies in the WBs because audio recordings are available in PE1 (e.g. *Listen and write the words in the correct column.*), whereas they are not available in P2 (e.g. *Tick the pairs that have the same vowel sound. or In each line, one word has a different vowel sound. Underline it.*). As in the SB, pronunciation is better linked to vocabulary in PE1. In both PE1 and P2, activities focused on the aspect of spoken form are found mainly in the SB, while the aspect of written form is represented mainly in WBs. The activities include labelling pictures and crossword puzzles, which fulfil the aspect of form and meaning as well.

The aspect of collocations is represented by 9.5 per cent of vocabulary activities in PE1 and 8.2 per cent in P2. In PE1, activities focus on expressions with verbs (e.g. *go, play, make, watch, read*), while in P2, they focus on the comparison (e.g. *as heavy as lead*). Both

coursebooks also focus on language functions and common everyday situations. In P2, there are short dialogues, for example, in a shop, in a restaurant. In PE1, the focus is, for example, on expressing offers, opinions, health problems, accepting and rejecting an offer. These activities are in the *Everyday English* section in D, *Revision* and *Progress check*. The activities in PE1 are more sophisticated than in P2; the *Everyday English* section is more straightforward and user-friendly. First, Winston the cat and Robopet point out the expressions to the learners, followed by work with the expressions (e.g. adding words or expressions to the sentences), then the learners get ready for dialogues in *Get ready to speak*, practise the dialogues and, finally, they work on their own dialogue with a peer. In PE1, the emphasis on language functions and common everyday situations is thus more evident.

The aspect of associations is represented more in PE1 (7.1%) than in P2 (2.5%). In both coursebooks, there are activities to categorise words and match words to the words of opposite meanings. For example, in P2, learners distinguish between categories within the vocabulary related to animals, food and drinks, and geographical features. In PE1, it is a question of categorising animals and vocabulary related to the description of a person, matching goods and specific shops or furniture and rooms. In addition, there are activities focused on adjectives of the opposite meaning in both coursebooks. In P2, these activities are found in the WB. In PE1, they concern not only adjectives but also adverbs. Thus, the activities in PE1 cover more topics and more parts of speech.

The aspect of word parts is represented in P2 by 5.4 and in PE1 by 3.2 per cent of vocabulary activities. In both coursebooks, this aspect is fulfilled with activities focused on the plural of nouns, ordinal numbers and the formation of adverbs from adjectives. PE1 also includes an activity on forming nouns from the verb and P2 an activity on compound nouns.

The aspect of concept and referents is represented only in P2, namely by two jokes, the point of which lies in the identical pronunciation of words (e.g. *weak* and *week*). The aspect of constraints on use is not directly represented by any activities. However, in PE1, learners in the *Everyday English* section are advised how to express interest and make polite requests (*We can ask for something politely like this...*). The number of activities fulfilling the aspect of concept and referents and the aspect of constraints on use seems to be related to the language level (A1). According to Brown (2010), many researchers suggest that learners need to understand better how lexical items complement each other, and it can be stated that the analysed coursebooks reflect this.

Sections focused on developing study skills and learning strategies are integral to current coursebooks. In PE1, these are represented by the *Study tip!* section. Learners are always encouraged to look closely at the pictures to help them understand the text and learn words in pairs, such as adjectives and adjectives of the opposite meaning (e.g. *difficult* vs *easy*). In addition, *Study tip!* in PE1 WB advises learners to write the word in a sentence because the word is better remembered in context.

3.5.2 Research question 2

Does the selected current English language student's book for Year 7 offer more vocabulary activities and cover more aspects of vocabulary knowledge than that published in 2000?

Project Explore 2 and *Project 3* (2nd ed.) are intended for learners in Year 7 or the corresponding year of multi-year grammar school. However, as there are eight units in *Project 3* (P3) and six in *Project Explore 2* (PE2), which is more appropriate for a school year, three selected units from each of the above were analysed.

While in PE2, each unit includes revision, a page on English-speaking countries, *Project* and cross-curricular links, in P3 *Revision*, *Culture* and *Project* occur only after even units; therefore, only sections A–D were analysed in the SBs. Moreover, unlike *Project 2*, *Project 3* has *Working with words* sections.

In the selected units of P3, fifty-eight vocabulary activities were identified, while in PE2, the number of vocabulary activities is higher (103). The category of use is the most frequently represented category in both SBs, followed by that of meaning; the category of form is the least represented of all. The difference lies mainly in vocabulary activities focused on the aspect of spoken form. In P3, the pronunciation section is at the end of the student's book, while activities focused on the aspect of spoken form in PE2 occur regularly in sections A–D.

Student's book	<i>Project 3</i> (2nd ed.)		<i>Project Explore 2</i>	
	T 58	%	T 103	%
Form	1	1.7	25	24.3
Spoken form	0	0	23	22.3
Written form	1	1.7	1	1.0
Word parts	0	0	1	1.0
Meaning	22	37.9	29	28.2

Form and meaning	19	32.8	25	24.3
Concept and referents	0	0	0	0
Associations	3	5.2	4	3.9
Use	35	60.4	49	47.6
Grammatical functions	12	20.7	23	22.3
Collocations	23	39.7	17	16.5
Constraints on use	0	0	9	8.7

SB – student’s book, WB – workbook, T – number of vocabulary activities in total

Table 3: Comparison of vocabulary knowledge in *Project 3* (2nd ed.) and *Project Explore 2*

Table 3 shows the number and percentage of vocabulary activities representing each of the three categories and each of the nine aspects of word knowledge according to Nation (2001). The results for the individual aspects will now be presented based on the percentage of vocabulary activities representing the aspects.

The aspect of form and meaning is the most frequently represented aspect both in PE2 (24.3%) and P3 (32.8%). Nevertheless, the number of activities is higher in PE2 (25) than in P3 (19). The current SB emphasises the multi-sensory approach, work with mixed-ability classes and learners with special educational needs. The already acquired vocabulary is activated through the *You first* section (e.g. *What’s your favourite city? Why? Describe it to your partner.*), followed by a picture (e.g. city map), and learners have to name places on the city map they already know (*Look at the map. Which places in the city can you name?*). Then learners name the places on the map using the words in the box (*Match the words in the box with the places on the map.*). Although both SBs work with the same city (London) and the real places in the city, P3 focuses more on the geographical names (e.g. *Nelson’s Column, Trafalgar Square*), whereas PE2 works with general terms first (e.g. *cathedral, bridge, square*).

The aspect of collocations is represented more in P3 (39.7%) than in PE2 (16.5%). Both SBs contain expressions within language functions and common everyday situations. For example, P2 focuses on making arrangements, giving advice and PE1 on expressing sympathy, regret, asking for and giving directions, persuading and encouraging. These activities are located in P3 in the *Everyday expressions* section and in PE2 in the *Everyday English* section. The expressions are presented in the peers’ photo stories in both SBs, followed by work with the expressions and preparation for dialogues. These activities then

fulfil the aspect of collocations as well as the aspect of grammatical functions. The activities in PE2 are more sophisticated than in P3; the *Everyday English* section is more straightforward and user-friendly. Learners get ready for dialogues following a scheme in *Get ready to speak*, practise the dialogues, and work on their own dialogue with their peers. In PE2, the emphasis on language functions and common everyday situations is thus more evident. In P3, more activities focus on expressions with verbs (*Match the verbs in A with the items in B.*) and phrasal verbs (*Match the words to make phrasal verbs from the story.*), especially in the *Working with words* section.

The aspect of grammatical functions is represented in the selected units of PE2 by 22.3 per cent and in P3 by 20.7 per cent of vocabulary activities. Learners use vocabulary in short dialogues (e.g. *Ask and answer questions about the places.*) and in short speaking and writing activities in both SBs. These also include dialogues in the *Everyday expressions* section in P3 and *Everyday English* section in PE2 (e.g. *Persuade your family and friends to try out interesting new things. Give good reasons! Make short dialogues.*), which fulfil the aspect of grammatical functions as well as the aspect of collocations.

The aspect of spoken form is represented only in PE2 (22.3%). In P3, the aspect is not represented since *Pronunciation* sections are not in the analysed sections of units. In PE2, pronunciation – as in PE1 – is directly part of the units (e.g. *Listen and check, then repeat.*). The aspect of written form is represented in both SBs by only one activity (e.g. *Write the correct weather*). However, activities focused on the aspect of written form aspect can be found in the workbooks.

The aspect of associations is represented in PE2 by 3.9 and in P2 by 5.2 per cent of vocabulary activities. Both SBs contain activities aimed at categorising words (e.g. furniture items and rooms in P3, places or institutions and a city/town in PE2) and matching words to the words of the opposite meaning (*Look at the words in A and match them to their opposites in B.* in PE2 and *Find the opposites...* in P3).

The aspect of constraints on use is represented only in PE2 (8.8%). These activities focus on colloquial language and can be found in the *Spoken English* section. The aspect of word parts is also represented only in PE2 (1.0%). It is an activity focused on the formation of adverbs from adjectives. Finally, the aspect of concept and referents is not represented in the analysed sections of units at all, which is related to the language level (A1–A2).

4 Conclusion

Based on the analysis of selected current EFL coursebooks and those published in 1999–2000, it can be concluded that current EFL coursebooks contain more vocabulary activities than those published twenty years ago. The former emphasise the multi-sensory approach, support mixed-ability classes and learners with special educational needs more. In both *Project 2* (1999) and *Project Explore 1* (2019), the aspect of form and meaning is the most represented; however, the number of activities is higher in *Project Explore 1*, where vocabulary work is more diverse than in *Project 2*. The aspect of grammatical functions comes second; the percentage is very similar, the difference being the higher number of activities in *Project Explore 1* workbook. On the one hand, the aspect of spoken form is more prominently represented in *Project 2*. On the other, pronunciation is better linked to the subject matter in *Project Explore 1*. While activities representing the aspect of spoken form occur in the student's books, the aspect of written form is represented primarily in the workbooks. Vocabulary activities representing the aspect of collocations in *Project 2* occur mainly in the student's book, while in *Project Explore 1*, the number of activities in the student's book and workbook is the same. The coursebooks do not contain activities fulfilling the aspect of constraints on use at all. Vocabulary activities fulfilling the aspect of concept and referents can be found only in a limited number in *Project 2*. Vocabulary in *Project Explore 1* is also part of the section on developing study skills and learning strategies.

Project Explore 2 (2019) includes more vocabulary activities than *Project 3* (2000). Compared to *Project 2*, *Project 3* emphasises the breadth of vocabulary knowledge in the *Working with words* section, which corresponds with the higher language level of the learners. In the selected units of *Project Explore 2* and *Project 3*, the aspects of collocations, form and meaning, and grammatical functions are covered most. More activities represent the aspect of collocations, and fewer activities represent the aspect of spoken form compared to *Project Explore 1* and *Project 2*, reflecting the higher level of *Project Explore 2* and *Project 3*. The aspect of constraints on use is represented only in *Project Explore 2*, reflecting the interest in spoken language in current coursebooks.

In *Project Explore 1* and *Project Explore 2*, I positively evaluate the work with vocabulary independently of grammar, emphasis on the multi-sensory approach and focus on everyday expressions and situations in the *Communication* and *Spoken English* sections. In *Project Explore 2*, I recommend including a section similar to *Working with words* in

Project 3 to support the depth of vocabulary, especially the aspect of word parts, associations and concept and referents.

The research aimed to verify the research tool, i.e. the categorical system based on Nation's (2001) nine aspects of word knowledge. The activities fulfilling individual aspects of vocabulary knowledge were specified and supplemented with examples of specific activities. The categorical system was verified through the agreement of two raters. It will now be used in the subsequent research, focusing on analysing two coursebooks, each consisting of four levels.

Notes

¹ <https://www.oxfordlearnersdictionaries.com/about/oxford3000> (2022). [online]

² <https://englishprofile.org/wordlists> (2022). [online]

³ <https://www.wgtn.ac.nz/lals/resources/academicwordlist> (2022). [online]

⁴ [https://16802fc1bf\(coe.int\)](https://16802fc1bf(coe.int)) (2022). [online]

⁵ In the Czech Republic, children attend elementary school from the age of six to fifteen. Stage 1 (Year 1–5) corresponds with primary school, Stage 2 (Year 6–9) corresponds with lower-secondary school. At the age of eleven, talented children can leave elementary school and go to lower-secondary grammar school.

⁶ https://elt.oup.com/catalogue/items/global/teenagers/project_explore/?cc=cz&selLanguage=cs&mode=hub (2022). [online]

References

- Bausch, K.-R., Christ, H. and Krumm, H.-J. (1997) *Handbuch Fremdsprachenunterricht*. Tübingen: Francke.
- Brown, D. (2010) 'What aspects of vocabulary knowledge do textbooks give attention to?' *Language Teaching Research* 15(1), 83–97. Online document. Accessed 30 May 2019. <https://doi.org/10.1177/1362168810383345>
- Council of Europe. *Common European Framework of Reference for Languages*. Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR) (coe.int)
- Cronbach, L. J. (1942) 'An analysis of techniques for diagnostic vocabulary testing.' *The Journal of Educational Research* 36, 206–217. <https://doi.org/10.1080/00220671.1942.10881160>
- Cunningsworth, A. (1995) *Choosing Your Coursebook*. Oxford: Macmillan Education.
- Dang, T. N. Y. and Webb, S. (2016) 'Making an essential word list for beginners.' In: Nation, I. S. P. (ed.) *Making and Using Word Lists for Language Learning and Testing*. Amsterdam: John Benjamins. 153–167. <https://doi.org/10.1075/z.208.15ch15>
- English Profile. *English Vocabulary Profile*. <https://englishprofile.org/wordlists>
- Falla, T. and Davies, P. (2007) *Maturita Solutions Pre-Intermediate*. Oxford: Oxford University Press.
- Falla, T. and Davies, P. (2008) *Maturita Solutions Intermediate*. Oxford: Oxford University Press.
- Fernández Orío, S. (2014) 'Vocabulary knowledge dimensions in EFL textbooks.' *Encuentro* 23, 30–37. Online document. Accessed 4 June 2019. <https://core.ac.uk/download/pdf/58911078.pdf>
- Haß, F. (2010) 'Methoden im Fremdsprachenunterricht.' In: Hallet, W. and Königs, F. G. (eds) *Handbuch Fremdsprachendidaktik*. Seelze-Velber: Klett/Kallmeyer. 151–156.
- Hedge, T. (2000) *Teaching and Learning in the Language Classroom*. Oxford: Oxford University Press.
- Janíková, V. (2005) *Osvojování cizojazyčné slovní zásoby: pedagogická, psychologická, lingvistická a didaktická východiska na příkladu němčiny jako cizího jazyka*. Brno: Masarykova univerzita.
- Jiménez Catalán, R. (2002) 'El concepto de competencia léxica en los estudios de aprendizaje y enseñanza de segundas lenguas.' *ATLANTIS* XXIV, 149–162.

- Lewis, M. (1993) *The Lexical Approach: The State of ELT and The Way Forward*. Hove: Language Teaching Publications.
- Lewis, M. (1997) *Implementing the Lexical Approach: Putting Theory into Practice*. Hove: Language Teaching Publications.
- Lund, R. (2006) 'Questions of culture and context in English language textbooks: A study of textbooks for the teaching of English in Norway.' Bergen: Universitetet i Bergen. Online document. Accessed 17 July 2022. <https://core.ac.uk/download/pdf/30823046.pdf>
- Meara, P. M. (2014) 'Vocabulary research in the Modern Language Journal: A bibliometric analysis.' *Vocabulary Learning and Instruction* 3(1), 1–28. Online document. Accessed 17 July 2022. <http://vli-journal.org/issues/onlinefirst/vli.v03.1.meara.pdf>
<https://doi.org/10.7820/vli.v03.1.meara>
- Milton, J. and Fitzpatrick, T. (2014) *Dimensions of Vocabulary Knowledge*. London: Palgrave Macmillan. <https://doi.org/10.1007/978-1-137-36831-7>
- Morgan, J. and Rinvoluceri, M. (2004) *Vocabulary*. Oxford: Oxford University Press.
- Nation, I. S. P. (1990) *Teaching and Learning Vocabulary*. Boston: Heinle, Cengage Learning.
- Nation, I. S. P. (2001) *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Nation, I. S. P. (2013) *Learning Vocabulary in Another Language*. (2nd ed.) Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139858656>
- Nation, P. and Webb, S. (2011) *Researching and Analysing Vocabulary*. Boston: Heinle, Cengage Learning.
- Oxford Learner's Dictionaries. *Oxford 3000*.
<https://www.oxfordlearnersdictionaries.com/about/oxford3000>
- Oxford University Press. *Project Explore*.
https://elt.oup.com/catalogue/items/global/teenagers/project_explore/?cc=cz&selLanguage=cs&mode=hub
- Richards, J. C. (1976) 'The role of vocabulary teaching.' *TESOL Quarterly* 10, 77–89.
<https://doi.org/10.2307/3585941>
- Römer, U. (2005) *Progressives, Patterns, Pedagogy: A Corpus-driven Approach to English Progressive Forms, Functions, Contexts, and Didactics. Studies in Corpus Linguistics*. Philadelphia: John Benjamins. <https://doi.org/10.1075/scl.18>
- Schmitt, N. and McCarthy, M. (1997) *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.
- Scrivener, J. (1994) *Learning Teaching*. Oxford: Heinemann English Language Teaching.
- Scrivener, J. (2011) *Learning Teaching*. London: Macmillan Education.
- Schmidtová, V. (2017) *Srovnávací analýza učebnic anglického jazyka pro 2. stupeň ZŠ z hlediska naplnění požadavků RVP ZV*. Online document. Accessed 17 July 2022.
<https://is.muni.cz/th/n7khf/schmidtova-dp.pdf>
- School of Linguistics and Applied Language Studies. *Academic Word List*.
<https://www.wgtn.ac.nz/lals/resources/academicwordlist>
- Sikorová, Z. (2010) *Učitel a učebnice: užívání učebnic na 2. stupni základních škol*. Ostrava: Ostravská univerzita v Ostravě. Pedagogická fakulta.
- Spratt, M., Pulverness, A. and Williams, M. (2011) *The TKT course: Module 1, 2 and 3*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139062398>
- Trnová, M. (2017) 'Obsahová analýza učebnic anglického jazyka pro 2. st. ZŠ se zaměřením na slovní zásobu na úrovni A2.' In: Janíková, V. and Hanušová, S. (eds) *Lingvisticky orientovaný výzkum v didaktice cizích jazyků*. Brno: Masarykova univerzita. 103–125.
- Trnová, M. (2021) 'Aspekty znalosti slovní zásoby v učebnicích anglického jazyka pro 2. stupeň ZŠ: Pilotní výzkum.' In: Janíková, V. and Hanušová, S. (eds) *Výzkum v didaktice cizích jazyků IV*. Brno: Masarykova univerzita. 137–172. <https://doi.org/10.5817/CZ.MUNI.P280-0055-2021-6>
- Ur, P. (2012) *A Course in English Language Teaching*. Cambridge: Cambridge University Press.
<https://doi.org/10.1017/9781009024518>
- Webb, S. and Nation, P. (2017) *How Vocabulary is Learned*. Oxford: Oxford University Press.
- West, M. (1953) *A General Service List of English Words*. London: Longman.

Sources

- Hutchinson, T. (1999a) *Project 2: Pracovní sešit*. Oxford: Oxford University Press.
Hutchinson, T. (1999b) *Project 2: Student's book*. Oxford: Oxford University Press.
Hutchinson, T. (2000) *Project 3: Student's book*. Oxford: Oxford University Press.
Phillips, S., Shipton, P. and Trnová, M. (2019) *Project Explore 1: Pracovní sešit*. Oxford: Oxford University Press.
Phillips, S. and Shipton, P. (2019) *Project Explore 1: Učebnice*. Oxford: Oxford University Press.
Wheeldon, S. and Shipton, P. (2019) *Project Explore 2: Učebnice*. Oxford: Oxford University Press.

Michaela Trnová teaches English and German at the Department of English Language Teaching at the Faculty of Education, University of Ostrava, Czech Republic. She specializes in teaching English as a foreign language, ELT coursebooks, materials and curriculum. She is a Ph.D. student of Didactics of Foreign Language at the Department of English Language and Literature at the Faculty of Education, Masaryk University, Brno, Czech Republic. She is currently working on her dissertation, which deals with the aspects of vocabulary knowledge in EFL coursebooks.

Address: Michaela Trnová, Department of English Language and Literature, Faculty of Education, Masaryk University, Poříčí 9, Brno 603 00, Czech Republic. [e-mail: michaela.trnova@osu.cz]

**Patterns and Variation
in English Language Discourse
9th Brno Conference on Linguistics Studies in English**

Mgr. Irena Hůlková, Ph.D., doc. PhDr. Renata Povolná, Ph.D., Mgr. Radek Vogel, Ph.D. (eds)

Published by Masaryk University Press,
Žerotínovo náměstí 617/9, 601 77 Brno, CZ
1st electronic edition, 2022

ISBN 978-80-280-0212-1

