The lexicon of the environment and its chemistry-related terms in ordinary discourse: Using social networks as corpora

Thesis

submitted and publicly defended on November 14, 2023

for the award of

Doctor of Philosophy of Université de Lorraine

Language sciences (*Sciences du langage*)

by

Tomara Gotkova

Committee:

Mathieu Constant – Chair

Rute Costa – Reporting examiner

Patrick Drouin – Reporting examiner

Lucie Barque – Examiner

Alain Polguère – Supervisor

Francesca Ingrosso – Co-supervisor

École doctorale Sociétés, Langages, Temps, Connaissances (SLTC)

To my parents, Irina and Yevgeniy,
and my grandfather Nikolay.

The earth is not a mere fragment of dead history,
stratum upon stratum like the leaves of a book,
to be studied by geologists and antiquaries chiefly,
but living poetry like the leaves of a tree,
which precede flowers and fruit,—not a fossil earth,
but a living earth; compared with whose great central life
all animal and vegetable life is merely parasitic.

Walden, 1854

Henry David Thoreau

Table of contents

Page

Chapter Page

Chapter Page

List of figures

Figure Page

List of tables

Table Page

Abstract

The main objective of our interdisciplinary research is to study the use of English environmental lexicon in ordinary discourse in the context of current and emerging environmental issues. First, we provide a sociological overview of environmental debate and the role of general public in climate change mitigation. We then formalize our vision of the terminology of current and emerging environmental issues and create a list of relevant environmental lexicon using the combination of Natural Language Processing techniques and manual selection.

By integrating a Natural Language Processing approach into the study of terminology, we contrast the notions of *keyword* and *Term*. Keywords represent semantically ambiguous linguistic forms used to build corpora of public environmental discussions from two social networks, *Twitter* and *Reddit*. By contrast, Terms are semantically disambiguated specialized lexical units. Furthermore, we present a multifaceted profiling of an environmental buzzword *carbon* and associated Terms. Such profiling is rooted in two interrelated tasks: (i) a linguistic study of *carbon* in specialized discourse; (ii) a comprehensive lexicographic description of *carbon* and related idioms. The findings reveal the complexity of *carbon*-related terminology and highlight the need for normalization.

To explore public understanding of *carbon*, we study the data from the social network corpora. We concentrate on colloquial interpretations of *carbon*, as opposed to its usage in specialized discourse, through the prism of *domestication* of Terms by the general language. For this, we propose a novel typology of terminological units based on their semantics and register.

Finally, we propose a list of terminological recommendations to address existing challenges of environmental terminology and its communication to the general public.

Résumé

L'objectif principal de notre recherche interdisciplinaire est l'étude du vocabulaire de l'environnement au sein du discours ordinaire dans le contexte des enjeux environnementaux actuels et émergents. Tout d'abord, nous présentons un aperçu sociologique du débat environnemental et du rôle du grand public dans l'atténuation du changement climatique. Ensuite, nous formalisons notre conception de la terminologie liée aux enjeux environnementaux actuels et émergents et établissons une liste du vocabulaire de l'environnement pertinent. Pour cela, nous combinons des techniques de traitement automatique des langues avec une sélection manuelle.

En intégrant l'approche du traitement automatique des langues à l'étude de la terminologie, nous comparons les notions de *mot-clé* et de *Terme*. Les mots-clés représentent des formes linguistiques sémantiquement ambiguës utilisées pour construire des corpus de discussions environnementales publiques extraits de deux réseaux sociaux : *Twitter* et *Reddit*. Inversement, les Termes sont des unités lexicales spécialisées et sémantiquement désambiguïsées. En outre, nous présentons une analyse approfondie du mot emblématique dans le domaine de l'environnement, *carbon* (**Fr.** carbone), ainsi que des Termes associés. Cette analyse repose sur deux aspects interreliés : (i) une étude linguistique de *carbon* dans le discours spécialisé ; (ii) une description lexicographique détaillée de *carbon* et des locutions liées. Les résultats révèlent la complexité de la terminologie autour de *carbon* et soulignent la nécessité de standardisation.

Pour explorer la compréhension publique de *carbon*, nous examinons les données issues des corpus de réseaux sociaux. Nous nous concentrons sur les interprétations non terminologiques de *carbon*, par opposition à son usage dans le discours spécialisé, à travers le prisme de la *domestication* des Termes par la langue générale. Dans cette optique, nous proposons une nouvelle typologie des unités spécialisées en fonction de leur sémantisme et du registre de langue auquel elles

appartiennent.

Finalement, nous formulons une liste des recommandations terminologiques visant à traiter les défis actuels de la terminologie environnementale et sa communication au grand public.

Acknowledgements

Throughout my Ph.D. journey, numerous colleagues and friends have made valuable contributions to my project. I would like to express my special thanks to:

my supervisor Alain Polguère for his professional guidance, patience, and exceptional linguistic intuition that always inspired me;

my co-supervisor Francesca Ingrosso for making this interdisciplinary project possible and for productive and engaging working meetings;

the members of my progress review committee, Michel Musiol and Manuel Ruiz-Lopez, for their valuable comments;

the personnel of the OLKi project, ATILF laboratory and the University of Lorraine for the financial and administrative support;

the colleagues in ATILF, Éva Buchi, Sandrine Ollinger, Veronika Lux Pogodalla, Nabil Gader, Yan Greub, Cyril Pestel, Olivier Servas, Laurent Gobert and Fabienne Dumont for their help and support;

the colleagues from the Pompeu Fabra University, Leo Wanner, Juan Soler Company and Alexander Shvets, for fruitful interdisciplinary collaboration;

the colleagues from the LEC-ru team, Svetlana Krylosova, Lidia Kolzun, Elena Akborisova, Polina Mikhel and Nikolay Chepurnykh for linguistic brainstorms;

our dear Igor Aleksandrovich Mel'čuk and Lidija Nikolaevna Iordanskaja for being our linguistic role models;

my friends Polina, Mariia, Anna and Mary for their endless support and encouraging conversations;

my parents who never stop believing in me, and my grandfather who would have been proud of me.

Abbreviations, symbols and writing conventions

Abbreviations

C	:	carbon
CCS	:	carbon (dioxide) capture and storage
CH_4	:	methane
CO_2	:	carbon dioxide
COCA	:	Corpus of Contemporary American English
En., **Fr.**	:	English, French
IPCC	:	Intergovernmental Panel on Climate Change
NLP	:	Natural Language Processing
OED	:	Oxford English Dictionary
UNFCC	:	United Nations Framework Convention on Climate Change

Symbols

'...'	:	signified – e.g. 'gas that contains carbon atoms'
⌜...⌝	:	idiom – e.g. ⌜CARBON DIOXIDE ⌝
X, Y	:	actant slots – e.g. *gas released by X*

Writing conventions

notion	:	important notion
VOCABLE$_{\text{(inherent gram. feat.)}}$	:	vocable name e.g. CLEAN$_{\text{(Adj)}}$
LEXICAL UNIT$_{\text{(gram. feat.)}}$ **no.**	:	lexical unit name e.g. CLEAN$_{\text{(Adj)}}$ **1**
signifier	:	e.g. *clean*
`Lexical function`	:	e.g. `Syn`, `Gener`, ...
`Lexical function`(*argument*)	:	lexical function application `Gener`(*carbon* **I.1**)
usage note	:	e.g. **spec**

Chapter 1

Introduction

Summary

1.1 Research context and objectives

Our research is focused on studying environmental terminology, specifically its usage in public discourse. The environment as a subject has gained social significance since the 1960s, leading to the proliferation of environmental terminology in everyday language. The widespread use of environmental expressions requires a comprehensive analysis of how these expressions are employed and interpreted by the general public.

In 1890s, Swedish scientist Svane Arrhenius made an important discovery by quantifying the effects of fluctuations in concentrations of atmospheric carbon dioxide on the Earth's surface temperature. He predicted that industrial emissions of carbon dioxide (CO_2) could potentially cause undesirable changes in global temperature in the future (Rodhe, Charlson and Crawford 1997). Over a century later, in the early 2000s, the scientific community reached a consensus regarding the rise in Earth's temperature due to man-made greenhouse gas emissions that contribute to the phenomenon known as *climate change*.[1] Given serious implications of climate change for humanity and planet's future, environmental topic has taken a firm place in the global agenda.

In response to environmental challenges, society is actively seeking mitigation solutions to minimize negative environmental impacts. These efforts embody the era of sustainability. The notion of environmental *sustainability* allows for multiple interpretations, but it commonly implies the harmony between the society and the natural world. In Miller and Spoolman (2014), sustainability is defined as follows:

> Sustainability is the ability of the earth's various natural systems and human cultural systems and economies to survive and adapt to changing environmental conditions indefinitely.

In an *environmentally sustainable society*, individuals are committed to sustain the environment and its natural resources, ensuring their availability for fu-

[1] The list of organizations adhering to this scientific consensus can be found at: https://climate.nasa.gov/scientific-consensus/.

ture generations (Miller and Spoolman 2014). Adhering to sustainable principles is currently associated with being a responsible member of society.[2]

In a recent international poll across 30 countries on public opinion on climate change, the respondents ranked climate change among the top 10 concerns, although it was considered less important than personal health, terrorism, educational prospects, and crime (Ipsos 2022). Another international poll conducted on Facebook showed that Europeans encounter discussions about climate change at least once per week through TV, news media, social media, personal conversations (Leiserowitz, Carman, *et al.* 2022). However, we believe that exposure to environmental Terms[3] is even more prevalent in the lives of ordinary citizens if we consider more subtle forms of encounters, such as:

- Food labels: *climate score A*;
- Travel tickets: *carbon footprint of the journey*;
- Banners: *protect the ecosystem*;
- Packaging: *biodegradable plastic*;
- Public rentals: *drive green*;
- Gadget use: *power consumption and carbon emissions*;
- Waste management: '*reduce, reuse, recycle*';
- Hotel practices: *eco-friendly spa treatment.*

The question is whether we use and interpret such environmental terminology in conformity with its scientific standards. Electronic dissemination of information has contributed to widespread assimilation of scientific knowledge among the general public. For example, in medical context, we observe the emergence of the concept of *e-patient* who possesses a relatively high level of medical literacy and the ability to use medical Terms knowledgeably as a result of exposure to online information (Láinez Ramos-Bossini and Tercedor Sánchez 2020). However, online information often lacks reliability, as mass media can provide biased, erroneous, and one-sided information. This is particularly pertinent to such topical domains

[2]https://www.pewresearch.org/global/2022/11/16/what-makes-someone-a-good-member-of-society/

[3]Term, with a capital *T*, refers to a specialized lexical unit, as opposed to a general language lexical unit. We introduce the notion in detail in 3.1.2.

as the environment. As asserted in Schaefer (2012: 533), "online media and blogs paint a picture of climate change that deviates significantly from the scientific view."

We analyze the interaction of the general public with environmental topic and its terminology through the prism of *domestication* that refers to the migration of Terms from specialized lexicon to the general language lexicon. Specifically, we analyze texts from two large social networks, *Twitter* and *Reddit*, to study how the reinterpretation of scientific concepts by the general public gives rise to alternative perceptions of these concepts. We are particularly interested in instances of uninformed usage of the environmental buzzword *carbon* and related expressions, considering that the current environmental discourse is centered around the need to reduce global carbon emissions.

> In July 2023, it was announced that Twitter changes its name to *X*. This poses a terminological problem for us since our study of Twitter concerns its 2020 version, prior to the major rebranding that took place in 2023. Therefore, throughout this text, we will continue to use the former name *Twitter* to refer to all the knowledge we accumulated before the name change. However, in cases where we need to discuss the current state of the platform, we will use the name *X*.

It is important to note that the environmental topic exists in a highly polarized and politicized context. Hence, we would like to make it clear that it is not in our present agenda to advocate for any social or political environment-related movement. Similarly, we do not support one single scientific point of view on existing environmental issues. Instead, we examine the environmental topic through the prism of linguistic analysis, which does not require taking a specific stance and allows us to remain as much as possible neutral.

> Nevertheless, we do acknowledge the scientific consensus on the reality of climate change as our default perspective on the ongoing environmental issues.

With the environment as underlying and principal theme of our research, our research objectives are rooted in terminology, lexicography, lexicology and qualitative corpus studies. Specifically, we focus on achieving the following objectives:

1. Examine the significance of environmental domain and its terminology within social context.
2. Propose the definition for the multifaceted concept of *the environment* and establish a respective thematic scope of our research.
3. Based on the established scope, compile a list of environmental vocabulary.
4. Develop theoretical and practical framework for building environment-focused social network corpora.
5. Analyze the integration of environmental vocabulary in the ordinary discourse through the prism of its popularization, with a specific emphasis on deviant usage.
6. Assess the potential of using social network texts for understanding everyday environmental language.
7. Propose a comprehensive lexicographic description of selected environmental vocabulary.

1.2 Related research

A substantial body of scientific literature has addressed various aspects related to our research, whether terminological, lexicographic or lexicological. However, to our knowledge, no prior research has synthesized these aspects, let alone having access to a collection of such works. Therefore, in order to narrow the focus of our literature review, we have chosen to prioritize social research on environmental domain and its terminology, with a particular emphasis on social media.

In 2000, the journal *Public Understanding of Science* featured a special review dedicated to social research on climate change (Trumbo and Shanahan 2000). The authors of this review discussed public literacy on climate change and the challenges associated with environmental communication. Following this work, the journal continued to publish articles examining public understanding of environmental science in almost every issue, with many of these articles relying on data

from social networks. In this regard, we will introduce two foundational notions:

> *Computer-mediated communication* (CMC) is a form of human interaction by means of computer-based tools, e.g., email, social networks, messengers, blogs, etc. (Bieswanger 2013).

> *Social media communication* is a type of computer-mediated communication that takes place on social networks or online platforms.

There is a consensus among social scientists on the use of online data, including information from social networks, websites, blogs, and comments on online articles as a useful (though occasionally controversial) source for studying environmental topic within social context. The significance of this data lies in its accessibility, richness, and the insights it provides into unfiltered and "authentic" public opinions and perceptions. Furthermore, this data serves as "real-world" material for linguistic analysis. The richness of this data stems from the diverse online community, given that social networks attract people of different ages, origins, educational and professional backgrounds, and social classes. Moreover, such platforms are characterized by the absence of formal restrictions on how one should express their[4] opinion, both in terms of content and linguistic form. Consequently, online public discourse is often informal, which makes it a valuable source for comparison with formal and strict specialized discourse.

In order to capture a general picture of public environmental debate and explore predominant concerns, views and sentiments, researchers aim to identify key topics, thematic communities and echo-chambers that shape them, as seen in Barr (2011), Cody *et al.* (2015), Williams *et al.* (2015), Elgesem, Steskal and Diakopoulos (2015), Maynard and Bontcheva (2015), and Veltri and Atanasova (2017). For instance, the article by Barr (2011) discussed three dominant public concerns identified in comments section of the *Guardian* articles: climate change skepticism, polarized views of the notion of sustainability, and critique of expert sustainable claims. In Elgesem, Steskal and Diakopoulos (2015), the authors

[4]Throught the text, we use the generic pronoun *they* to remain gender-neutral when referring to an unspecified individual.

adopted a broad perspective of the English Web blogosphere related to climate change analyzing 3,000 blogs and concluded with predominant attitudes being neutral, skeptical and accepting, and with predominant topics being climate science and politics.

Moving from a broader context to more specific subjects, social media, especially social networks, provide researchers with a unique advantage to access real-time content or time-specific archived data. Driven by journalistic interest to firsthand experiences, researchers gain access to online discussions related to specific events such as natural disasters, thematic conferences, protests, and press releases (Segerberg and Bennett 2011; Pearce, Holmberg, *et al.* 2014; Newman 2017; Roxburgh *et al.* 2019). For example, Pearce, Holmberg, *et al.* (2014) used Twitter data to identify prominent hashtags and thematic groups that formed after the release of the *2013 IPCC Working Group 1 Report*. Roxburgh *et al.* (2019) studied public response on Twitter to extreme weather events – Hurricane Irene (2011), Hurricane Sandy (2012) and Snowstorm Jonas (2016), – concluding that discourses varied across the three event accounts.

To highlight the intersection of *expert* and *public* perspectives, researchers often explore *public understanding*, *public perception* and *public acceptance* of current environmental problems and emerging solutions to these problems, such as research on public response to renewable energy, carbon management technologies, and climate mitigation policies (Wallquist, Visschers and Siegrist 2009; Boudet 2019; Simonchuk and Romasheva 2021; S. Y. Kim *et al.* 2021; Dabla-Norris *et al.* 2023). For instance, Wallquist, Visschers and Siegrist (2009) investigated "lay concepts" of carbon dioxide capture and storage technology through qualitative interviews. S. Y. Kim *et al.* (2021) assessed public acceptance of solar energy on Twitter.

Additionally, public perceptions of environmental topics are also studied through surveys and interviews of various scale, ranging from small-scale studies (Henry 2000; Sterman and Sweeney 2002) to national (Leiserowitz, N. Smith and Marlon 2010) and international (Lorenzoni *et al.* 2006) surveys.

To account for other actors participating in online environmental discussions, current research extends beyond public opinions and addresses such actors as ex-

perts, journalists and politicians, with specific attention to experts as bearers of environmental knowledge. However, a review by Schaefer (2012) found that climate experts were not highly active on social media, and individual scientists preferred interacting on thematic blogs. Walter, Lörcher and Brüggemann (2019) investigated networking patterns of scientists discussing climate change on Twitter and concluded that interactions were mainly with peers. Additionally, comparative research, as in Haunschild *et al.* (2019), examined differences in climate change topics of interest between the general public and researchers on Twitter.

While above-mentioned works mainly focus on factual knowledge, social and linguistic research also explores the interaction of the general public with environmental terminology. Specifically, researchers have pinpointed public understanding of environmental headline Terms, e.g., *biodiversity*, *eco-friendly*, and *sustainable* (Thornton 2009; Fletcher and Downing 2011; Campbell *et al.* 2015). The two competing Terms, *climate change* and *global warming*, received specific attention from researchers (Whitmarsh 2009; Jang and Hart 2015; Shi *et al.* 2020). Several studies highlighted the need to clarify the meaning of newly emerging or established environmental Terms (Glavič and Lukman 2007; Helbig *et al.* 2022).

Due to carbon-driven environmental agenda, environmental buzzword *carbon* is of particular interest. For example, context-dependent senses of *carbon* have been studied (Whitmarsh, Seyfang and O'Neill 2011; Bumpus 2011; Twyman, T. A. Smith and Arnall 2015; Loeve and Bensaude-Vincent 2017; Olfe-Kräutlein *et al.* 2022). Several studies revealed the need to clarify *carbon*-related terminology such as *carbon neutral* and *carbon footprint* (Wiedmann and Minx 2007; D. Pandey, Agrawal and J. S. Pandey 2011; Watkins and Durning 2012; Yu, Hoepner and Adamsson 2016; Trouwloon *et al.* 2023). Due to the intersection of various discourse genres within environmental topic, it is also important to study the migration of environmental vocabulary between them, as demonstrated with *carbon neutral* in the work of Dury (2008).

We would like to single out the researchers Nelya Koteyko[5] and Brigitte Nerlich[6] who have made a significant contribution to the study of *carbon*-related ter-

[5]https://www.qmul.ac.uk/sllf/language-centre/people/academic/profiles/koteyko.html

[6]https://www.nottingham.ac.uk/sociology/people/brigitte.nerlich

minology from a linguistic perspective. Their extensive and mainly collaborative work addressed *carbon* phraseology through the prism of metaphors, frames and lexical creativity (Nerlich and Koteyko 2009; Koteyko 2010; Koteyko, Thelwall and Nerlich 2010; Nerlich and Koteyko 2010; Nerlich, V. Evans and Koteyko 2011; Nerlich 2012; Nerlich and Jaspal 2013; Nerlich and Hellsten 2014).

In the review of papers on climate change discussions on social media conducted by Pearce, Niederer, *et al.* (2019), the authors identified three points of research gap to be addressed: lack of qualitative studies, scarcity of research on visual communication, and unbalanced representation of social media with the majority of studies relying on Twitter. While we do not claim to address all these gaps, we do propose a qualitative analysis of Reddit as an alternative social media platform.

1.3 Theoretical and methodological framework

Environmental terminology. We study the environmental domain, and by extension its terminology, taking into consideration its following properties: (i) dynamism; (ii) interdisciplinarity; (iii) social implications. The environmental domain is dynamic as it constantly adapts to current scientific findings and technological advancements. Interdisciplinarity of the domain stems from the range of scientific disciplines which contribute to environmental knowledge. By studying social implications, we account for the crucial role of environmental terminology in environmental dialogue between experts and within the general public. Furthermore, we believe that the incorporation of environmental terminology into public discourse is not just the matter of vocabulary, but it determines public attitudes to current environmental issues and, more importantly, shapes public acceptance of mitigation policies.

We adopt a synchronic perspective and study the use of environmental vocabulary in the modern social network discourse. A significant challenge that arises while working with environmental terminology lies in its versatility and pertinence to various domains. To address this challenge, we defined the scope of our study which focuses on the *terminology of current and emerging global environmental issues*. Such issues include biodiversity, climate change, ecosystem degradation,

pollution, etc.

The objective of studying the terminology of current and emerging environmental issues is twofold. First, our goal is to identify what linguistic expressions shape this subject and compile a list of corresponding vocabulary. Second, we aim to perform a lexicographic analysis of selected environmental lexical units.

Explanatory and Combinatorial Lexicology. In our lexicographic analysis of environmental Terms, we follow the principles of the lexical-semantic approach. Specifically, we use the principles of Explanatory and Combinatorial Lexicology, which is the lexical module of Meaning-Text theory (Mel'čuk 1995). This theoretical framework provides us with a well-structured and comprehensive methodology for systematic and formalized analysis of any natural language lexicon. Our analysis relies on the two fundamental notions of *vocable* and *lexical unit*, which refer to the main units of lexicographic description, and which are discussed in detail in Polguère (2016):

> By analogy with traditional dictionaries, *vocable* refers to a dictionary article which is constituted of *lexical units*, i.e., senses. For example, the vocable NOTEBOOK contains two lexical units: NOTEBOOK **1** 'pad for notes' and NOTEBOOK **2** 'small computer'. These two lexical units account for the *polysemy* of the vocable NOTEBOOK.
>
> Lexical units can be further divided into two types: lexemes and idioms. A *lexeme* is a set of wordforms that are inflectional expressions of a given sense. For example, the verbal lexeme $\text{SING}_{(\text{V})}$ **1** is a set of wordforms and phrases such as *sing*, *sang*, *is singing* which all refer to the sense 'to use voice to create a melody'. An *idiom* is a semantically non-compositional phraseme. For example, the idiom ⌜BE ALL EARS⌝ refers to readiness to listen attentively, as in *I'm all ears*.

Using these notions, we model the lexicon with a specific emphasis on combinatorial properties of the lexical units. The result of such lexicographic description represents a lexical network of interconnected lexical units.

As mentioned earlier, the principles of Explanatory and Combinatorial Lexicology can be universally applied across different natural languages. Similarly, the lexical-semantic approach has proven to be effective for terminological analysis (Jousse and Bouveret 2003; Dancette and L'Homme 2004; Bonadonna 2020). By applying these principles to environmental terminology, we ensure a rigorous lexicographic analysis of environmental specialized lexical units from multiple perspectives, including the description of polysemy, combinatorics, definitions, and usage examples.

The theoretical approach of Explanatory and Combinatorial Lexicology is compatible with computational lexicography. We use the English Lexical Network as a machine-tractable lexical resource to implement our lexicographic analysis. The English Lexical Network developed at the ATILF CNRS laboratory is a non-ontological lexical network that models the lexicon in graph format (Polguère 2014). Additionally, it constitutes a part of a multilingual set of lexical networks, available in Russian, French and other languages.

Although the English Lexical Network is originally a general language lexicon system, it allows for the integration of specialized vocabulary. The same principles are applied to modeling general language lexical units and specialized lexical units. Therefore, both general language lexical units and specialized lexical units receive the same treatment (Ingrosso and Polguère 2015).

Finally, we advocate for descriptive lexicographic approach, as opposed to prescriptive one. This implies that we aim to document and account for the relevant senses present in the lexicon, regardless of their conformity with linguistic norms.

Domestication of Terms. We discuss the interaction of the general public with environmental lexicon in the context of *domestication* of Terms. *Domestication*, in its original sense used to describe the taming of wild animals and plants by humans, refers to the process of the public becoming familiar with specialized vocabulary and incorporating it in everyday discussions. This process can lead to either informed usage of Terms that conforms to terminological standards or uninformed usage, which may involve significant deviations from scientific concepts.

Domestication is particularly pertinent to the environmental terminology, where many environmental concepts leave the realm of "pure" environmental science and become "common knowledge."

Within domestication strategy to terminology, we propose the following typology of Terms based on their semantic properties and their pertinence to register: *full Terms* (technical Terms), *runaway Terms* (popularized Terms) and *quasi-Terms* (distorted lay interpretations of Terms). We are particularly interested in quasi-Terms. They are "distorted" instances of uninformed, incorrect, or biased usage of scientific terminology by non-experts. We provide a more detailed discussion of the typology in Chapter 3.

The domestication of Terms in general language can be attributed to various factors. One of the reasons for domestication of Terms is the interest of the general public in specific domains. Moreover, modern technologies have played a significant role in the widespread use of Terms in everyday language. With the abundance of domain-specific information available on the internet, people can easily access and acquire new knowledge, including new terminology. For instance, typical online news sources feature dedicated sections for various domains such as economics, health, environment, science, sports, and more. Other channels of knowledge dissemination include educational textbooks, popular science literature, encyclopaedias, etc. (Beacco *et al.* 2002). Personal experiences also contribute to the intergration of Terms into general language. For example, individuals dealing with a medical condition might use medical terminology related to their situation (names of illnesses, symptomps, medications, etc.) in their everyday conversations, regardless of their own or their interlocutor's expertise in the field.

Domestication falls within the scope of the linguistic phenomenon known as *determinologization* which consists in migration of specialized lexicon into general lexicon (Meyer and Mackintosh 2000; Halskov 2005; Nová 2018; Picton, Condamines and Humbert-Droz 2021). It is important to note that in literature on determinologization, it is metaphoric senses that are qualified as emblematic cases of determinologized Terms. In (Nová 2018: 394), the author discusses an example of the geology Term *atmosphere* 'a layer of gases that surrounds the planet' and its general language metaphoric counterpart *atmosphere* 'mood of a place or an

event'.

In our approach, we propose to differentiate between metaphoric senses and distorted senses, i.e., quasi-Terms. For example, the medical Term *tumor* refers to the swelling of a body tissue. Tumors can be either benign (non-cancerous) or malignant (cancerous). In general language, *tumor* is often associated with cancer with no further distinction between cancerous and non-cancerous tumors. Therefore, the general language sense of *tumor* 'cancerous swelling' is a quasi-Term which represents an erroneous and one-sided intepretation of a scientific concept.

Finally, due to the gradual nature of determinologization, it is commonly studied from a diachronic perspective, as seen in the works of Dury (2008), Nicolae and Delavigne (2009), and Humbert-Droz and Picton (2022). However, we adopt a synchrony perspective on domestication of Terms, focusing on the immediate context of the modern English.

1.4 Limitations

We acknowledge the five following limitations of the research which concern our expertise in the environmental domain, corpus compilation and usage, keyword list, and lexicographic analysis.

1. **Expertise in the environmental domain**. While our research discusses the environmental domain and environmental terminology, we are not the experts in this field. To overcome this limitation, we studied scientific and educational literature on the subject and actively followed related news. Furtemore, we benefited from a collaboration with Francesca Ingrosso, an expert in green chemistry,[7] who provided us with domain-specific knowledge, particularly clarifying specific technical aspects.
2. **Representativeness of the corpus**. A substantial part of our research addresses the compilation of social network corpora for the purpose of studying the usage of environmental lexicon in ordinary discourse. It is important to

[7]According to Manahan (2004), green chemistry is a "practice of chemical science and manufacturing in a manner that is sustainable, safe, and non-polluting and that consumes minimum amounts of materials and energy while producing little or no waste material."

note that our corpora, built on data from Twitter and Reddit, are not representative of the entirety of social media discourse due to their collection within a restricted timeframe and specific criteria. Similarly, our subcorpus sample is not fully representative of our larger Twitter and Reddit corpora, as it was compiled using specific keywords.

3. **Completeness of the keyword list**. One of the outcomes of our research is presented as a list of environmental keywords which can be further used for various terminological, pedagogical and Natural Language Processing tasks. This is a dynamic list which is neither exhaustive nor finite. On the contrary, we recognize the necessity of a constant update by adding new units that represent emerging environmental concepts and removing those that lose their relevance.
4. **Objectivity of qualitative analysis**. To study the usage of *carbon* in ordinary discourse, we employed the technique of qualitative content analysis (Schreier 2013) by performing close reading of texts in a subcorpus sample. While the objective of this analysis was to gain insights into the senses attributed to *carbon* by users, our analysis was limited to the immediate context of each text. We did not have access to extra-linguistic information such as the author's background or pragmatic intentions. Therefore, we acknowledge that certain texts allow for multiple interpretations in the absence of a broader context.
5. **Targeted lexicographic analysis**. We conducted lexicographic analysis of the following vocables related to the environmental topic: CARBON, ⌜CARBON DIOXIDE⌝, ⌜CARBON CYCLE⌝, ⌜CARBON FOOTPRINT⌝, and ⌜CARBON CAPTURE AND STORAGE⌝. Our primary goal was to develop environment- and chemistry-related senses of these vocables by providing detailed definitions, semantic links, illustrative examples, etc. For senses unrelated to our subject, we did not provide the same level of thorough treatment. Furthermore, we acknowledge the possibility of overlooking the inclusion of senses that may be relevant to these vocables but lie outside of our focus.

1.5 Structure of the thesis

The thesis is organized into six chapters. The present introductory Chapter 1 outlines the research context, objectives, methodology, and limitations.

In Chapter 2, we introduce the topic of the environment, examining it from both scientific and sociological perspectives. We define the notion of *the environment* by drafting the polysemy of the corresponding vocable ENVIRONMENT. Furthermore, we explain the difference between a broader notion of the environment as opposed to the narrower focus of the notion of *climate change*. We describe the significance of the modern environmental debate through the prism of environmental science and public engagement with environmental issues. Within the public engagement perspective, we address such subjects as environmental awareness, public efforts in environmental mitigation, characteritics of environmental communication, polarization of public attitudes and the level of public literacy regarding environmental knowledge.

In Chapter 3, we adopt a terminological perspective to address the environmental topic. First, we introduce the notions of specialized communication, Term and terminology. Furthermore, we discuss the theoretical aspects of a linguistic phenomenon known as *determinologization* and contrast it with our proposed notion of *domestication*. Following this discussion, we propose a novel classification of Terms into three distinct types: full Terms, runaway Terms and quasi-Terms. To conclude the chapter, we analyze the characteristics of environmental discourse and its terminology.

In Chapter 4, we describe the process of buidling social network corpora which consists of two major stages. The first stage involves the compilation of the list of environmental keywords using automated keyword extraction techniques. The second stage consists in using this keyword list to extract relevant thematic texts from two online platforms: Twitter and Reddit. We conclude the chapter by providing a quantitative summary of the extracted data for both corpora. Additionally, we discuss the similarties and differences between the two platforms.

In Chapter 5, we present a linguistic analysis of the environmental buzzword *carbon* to explore its usage in specialized discourse. Based on these analyses, we

offer a lexicographic description of the following vocables in the English Lexical Network: CARBON, ⌜CARBON DIOXIDE⌝, ⌜CARBON CYCLE⌝, ⌜CARBON FOOTPRINT⌝ and ⌜CARBON CAPTURE AND STORAGE⌝.

In Chapter 6, we describe a qualitative study of the *carbon* subcorpus composed of tweets and Redit posts. We aim to investigate how the general public perceive, conceptualize and lexicalize *carbon* in online discussions. As a result of our analysis, we present a list of identified conceptualizations.

Finally, in the concluding Chapter 7, we summarize the findings and outline possible directions for future research.

Chapter 2

The environment as a social, political and scientific construct

Summary

While our main objective is to study the lexicon of the environment, we consider it important to look at the broader environmental context that is deeply rooted in sociology. This is because language does not exist in a vacuum, and there are various cultural, social, economic, and political extralinguistic factors which eventually influence the way we talk about the environment. Examining these factors provides a more comprehensive understanding of the environmental discourse and its terminology.

In Section 2.1, we define the notion of *the environment* and the related notions, such as environmental issues, and climate change.

In Sections 2.2 and 2.3, we take a look at environmental science and the modern environmental movement, with a specific focus on how the general public engages with the environmental topic.

2.1 What does *the environment* mean?

2.1.1 Drafting the polysemy of the vocable ENVIRONMENT

One of the central notions of our research is the notion of *the environment*. To clarify what we mean when discussing the environment, environmental issues, environmental awareness, etc., we would like to explain what senses are associated with *environment* as a linguistic entity. For this, we will now draft the general polysemy of the vocable ENVIRONMENT. After conducting a lexicological study of the vocable ENVIRONMENT, we concluded that it comprises four senses:

- ENVIRONMENT **I** [Ex.: *mountain environment with trees*]
- ENVIRONMENT **II** [Ex.: *laws which protect the environment*]
- ENVIRONMENT **III** [Ex.: *pleasant workplace environment*]
- ENVIRONMENT **IV** [Ex.: *graphical programming environment*]

ENVIRONMENT **I** refers to an area which surrounds a person or an object in which they live or exist:

(1) Nearly 400,000 people were resettled but millions continue to live in an **environment** where continued residual exposure created a range of adverse effects.
[COCA:[1] Helen Caldicott, Nuclear Power Is Not The Answer, 2012, http://www.thirdworld traveler.com/Helen_Caldicott/NuclearPowerNotAnswer.html]

ENVIRONMENT **II** refers to the natural world and its interconnection with the modern society, as perceived through the prism of climate change and harmful human activity, see (2). This sense has gained prominence in the context of increasing environmental risks and it is usually used with the definite article. Furthemore, the use of *the environment* is restricted to the use in singular form only (*We must save the *environments*). In 2.1.2, we will present and formulate the definition of the notion of *the environment* as we use it for the purposes of our research.

(2) This ever-growing increase of chemical substances represents a primary issue for **the environment** and human safety.
[COCA: Gadaleta Dominico, SAR and QSAR modiling of a large collection of LD50 rat acute oral toxicity data, *Journal of Cheminformatics*, 2019.]

ENVIRONMENT **III** refers to a context in which someone exists, and which can have a positive or negative impact on their existence, see (3). Such context is defined by a set of conditions. For example in (3), *the best environment for children with handicaps* is presumably defined by home settings and spending time with family members which have beneficial effect on children.

(3) Most child-welfare workers now believe that family settings provide the best **environment** for children with handicaps.
[COCA: K. Springen, B. Kantrowitz, The long goodbye, *Newsweek*, 1990, Vol. 116 Issue 17, p77, 3p.]

ENVIRONMENT **IV** refers to a system used to operate a computer or computer-like device:

(4) The Zip format is the most popular compression format used in the Windows **environment**, and WinZip is the most popular compression utility.
[COCA: *WinZip – Basic Information*, 2012, http://www.winzip.com/aboutzip.htm.]

[1] *COCA* stands for the *Corpus of Contemporary American English*: https://www.english-corpora.org/coca/.

2.1.2 *The environment*: study of the notion

As follows from the previous section, ENVIRONMENT II, which is one of the senses of the vocable ENVIRONMENT, denotes the Earth's natural world that is being *destroyed*, *polluted*, *harmed* and even *killed* by human activity, which is why it has to be *protected*, *conserved* and *saved*. When we examine the definitions of this sense in current lexicographic sources, they usually refer to the Earth's natural world (5), with some exceptions when dictionaries include other elements, such as the human element (6). Despite the fact that the human element is not mentioned in definitions like (5), the provided usage examples are mostly related to damaging (7) and protecting (8) the environment.

(5) The natural world in which we live. [*Oxford Dictionary of Environment and Conservation*, 2013]

(6) The natural world or physical surroundings in general, either as a whole or within a particular geographical area, esp. as affected by human activity. [*Oxford English Dictionary*, online]

(7) Many modern farming methods are highly damaging to the environment. [*Longman Dictionary of Contemporary English*, online]

(8) We need to protect our environment for future generations. [*The American Heritage Dictionary of the English Language*, online]

While the definition (5) only refers to nature, such as fauna, flora, and other biological components, *the environment* has come to mean much more. However, despite its prominence in the modern environmental debate, there is no commonly agreed definition of *the environment* in the context of the current environmental challenges. Therefore, we will propose our own interpretation of the notion of *the environment* as it is incorporated in the modern environmental discourse and as it will be used throughout the text.

To understand what the environment means in environmental discourse, it is necessary to adopt a wider perspective. In modern times, the world is seen

as balancing on the brink of the upcoming environmental disaster, largely due to changes in the Earth's climate. According to scientific evidence, these changes are caused by both natural and human-related factors, with the human element being the dominant factor. Therefore, we believe that anthropocentrism should be a strong semantic component to consider when attempting to define *the environment*. Furthermore, as Lakoff (2010: 76) rightly said, "the environment is not just about the environment, [...] it is intimately tied up with other issue areas: economics, energy, food, health, trade, and security." Indeed, the interaction of humans with the natural environment has numerous social, political, and economical implications, which also substitute the essence of the environment. Furthermore, it is worth noting that a significant shift in the way scientists, activists, and the general public view the environment happened after World War II. Prior to this, environmental concerns were often focused on local issues. However, with the rise of global issues such as climate change, and with the development of technologies that allow us to study these issues on a global scale, there was a shift towards viewing the environment as a global rather than a local system (Praskievicz 2021). Finally, the idea of protecting the environment is closely associated with the future. One of the primary motivations for environmental mitigation is to preserve the environment for future generations.

> In conclusion, we use the notion of *the environment* to refer to the natural world of the Earth as it is perceived today through numerous environmental challenges caused by human activity. These challenges have severe consequences that will extend to future generations.

Terminological clarification. Now that we have defined *the environment*, we need to clarify its relationship with other related notions. In environmental discourse, notions such as *nature*, *ecology*, *the Earth*, *planet*, and *climate* are often used interchangeably with *the environment*, although *the environment* is a vaster notion that encompasses all of them. These notions are often used in discussions of risks related to the environment.

When we talk about *environmental issues*,[2] we are referring to a generic notion

[2]In literature, negative environmental changes are also referred to as *environmental concerns*,

that describes negative changes in the environment that are a serious concern. Although there is no official list of environmental issues, environmental organizations prioritize issues such as biodiversity loss, climate change, global warming, greenhouse effect, overpopulation, pollution, sea-level rise, and waste management, among others.

Climate change, which refers to "long-term shifts in temperatures and weather patterns" (UN 2023), has received particular attention from the media and scientists. There has been a shift in focus from the environment as a whole towards climate change as a specific phenomenon, which is considered the most critical environmental issue today, with far-reaching impacts on natural ecosystems and human society. As a result, the Terms *climate* and *climate change* have become increasingly prevalent in environmental discourse. This led to the emergence of notions related to climate change only, such as *climate* (*change*) *awareness*, *climate* (*change*) *literacy*, *climate* (*change*) *debate*, *climate* (*change*) *mitigation*, *climate* (*change*) *movement*, *climate* (*change*) *narrative*, etc. Furthermore, many news outlets have created separate sections for environmental news, categorizing them under "Environment" or "Climate."

While climate change is a critical issue, it is not directly related to other environmental issues, such as deforestation, water pollution, and land degradation. Despite its prominence, it is important not to reduce everything to the climate and climate change only.

> To avoid any terminological confusion, we will prioritize the Term *the environment* to refer to the notion discussed in the present section, and the Term *environmental issues* to refer to any problem associated with the deterioration of the environment. In cases when these designations can distort the message when we cite the literature, we will preserve the original terminology used in the source.

---

environmental problems, *environmental challenges*, etc.

2.2 Environmental science

Environmental science is a diverse and interdisciplinary field which aims to study and address the environmental issues that the modern world is facing nowadays both on global and local levels. Miller and Spoolman (2014: 7) hightlighted the following fields and subfields related to the environmental science: biology (ecology, botany, zoology), chemistry (biochemistry), earth science (climatology, geology, hydrology, paleontology), social sciences (anthropology, demography, geography, economics, political science), and humanities (history, ethics, philosophy). The collaborative efforts of environmental scientists, who bring knowledge from various fields, allow them to study the human impact on the environment effectively.

Environmental science is a relatively new field which began to form in the second half of the twentieth century as a response to the emergence of polluting technologies and, as a result, increased scientific and public concern about the environment. Environmental science was developing alongside with environmental movement, which at that time was manifesting itself in numerous forms, such as local activist campaigns and the foundation of small-scale groups and organizations aimed at drawing attention to environmental issues.

As a dymanic and evolving field, environmental science constantly adapts to the new technological advancements and emerging environmental issues. It primarily focuses on the interactions between humans and nature, the current and emerging environmental issues, and the potential solutions to such problems. The list of current environmental issues changes with time, as some issues disappear, while others emerge. For instance, Europe and the North America were facing the problem of acid rain in the 1970s and 1980s. Acid rain is a form of precipitation with a pH level harmful for vegetation, land and water surfaces. The regulations aimed at the reduction of emissions contributing to acid rains in USA (EPA 2022a) and Britain (McCarthy 2010) solved the problem. What used to be a big issue at that time is now a topic of marginal interest (on a global scale) only relevant to regions with poor emissions control.

Given that environmental issues affect the future of the Earth and humanity, environmental science has a significant societal and political impact. Scientific

findings on the current state of the environment provide the basis for environmental policies implemented by governments and companies worldwide. International environmental cooperation is based on scientific findings about anthropogenic climate change due to elevated greenhouse gas emissions, which formed the basis of numerous international agreements such as the Montreal Protocol (1987), UN Framework Convention on Climate Change (1992), Kyoto Protocol (1997), and Paris Agreement (2015). Governments commit to adjust local policies to curb greenhouse gas emissions according to these agreements.

Environmental science has been taught in schools since the 1970s. It was a part of school curriculum in all the states of the United States by the end of the 1970s (Masih 2017), and in Europe, it was legitimated by the Resolution of the Council on Environmental Education in 1988 (Hesselink and van Kempen 1999). At present, educational institutions across the globe offer a great variety of environment-related programs. Graduates in environmental science have excellent employment opportunities, as environmental specialists are in high demand in the modern world.

2.3 Public engagement with the environment

2.3.1 Global environmental awareness

The notion of *environmental awareness* refers to consciousness and knowledge about the environment, the related environmental issues, and empathy towards the topic. This phenomenon has now extended beyond academic problems and affects personal and social lives. The internet and other media channels have facilitated access to information about the environment, resulting in more people becoming aware of environmental issues. As a result, people can engage with environmental awareness on different levels, such as expressing concern, just being aware, or being aware and taking action.

The evolution of public engagement with the environmental topic has undergone different stages of development over time (Capstick *et al.* 2015). From 1980s to the mid-2000s, people became aware of the topic and started showing public concern. However, from the mid- to late-2000s, public concern declined, and

skepticism increased. By 2007, it was apparent that support for climate change mitigation had grown to a global scale, with a stabilization of the level of public concern in the 2010s.

It is interesting to observe the diversity in public attitudes towards environmental issues. Surveys conducted in several countries such as Australia, Germany, India, and the United States have identified various public categories based on their attitudes towards global warming. These categories included Alarmed, Informed, Concerned, Undecided, Experienced, Dismissive, Indifferent, Cautious, Disengaged, and Doubtful. The Cautious group was the largest in Germany and Australia, while the Concerned group was the biggest in the USA and the Experienced group in India (Morrison *et al.* 2013; Metag, Füchslin and Schäfer 2017; Leiserowitz, Roser-Renouf, *et al.* 2021)

Although there is no unanimous opinion as to what event gave birth to the modern environmental movement, the American scientist and writer Rachel Carson is considered one of its main inspirers. In 1962, she published the book *Silent Spring* which is considered a significant milestone in raising both scientific and public awareness about environmental issues. The book drew attention to the detrimental effects of agricultural pesticides on the environment and accused the chemical industry of dishonesty. This led to the banning of DDT[3] pesticides in the United States. It had significant environmental and social benefits, leading to reduction in environmental contamination, recovering of species, development of alternative pesticides and protecting public health. In 1970, the American senator Gaylord Nelson mobilized 20 millions of Americans for a nationwide street protest[4] against air and water pollution which went down in history as the first Earth Day. This movement has now become global and unites around one billion of environmental activists from more than 190 countries.[5] Around the same time, concerns arouse about the future of human civilization. In 1970, a report from the Massachusetts Institute of Technology predicted a potential civilization collapse due to the overuse of natural resources, population growth, and uncontrolled

[3]DDT stands for dichlorodiphenyltrichloroethane, which is a chemical compound.

[4]To this day, the 1970 Earth Day is the second largest protest in the USA.

[5]https://www.earthday.org/history/

pollution (Meadows *et al.* 1972). Clearly, the pop culture did not stand aside as environment-related movies, songs, and fiction books were being created to address environmental concerns. For example, the dystopian movie *Soylent Green* (1973), which depicted futuristic environmental disaster, was a big hit at that time.

Environmental disasters, such as the Bhopal (1984) and Chernobyl (1986) disasters, the Amoco Cadiz oil spill (1978) and the Exxon Valdez oil spill (1989), played a significant role in the development of environmental science and environmental movement. These events made society admit the need for serious local and international regulations to prevent further harm to the environment. This, among other things, led to the foundation of several large governmental and non-governmental environmental organizations, including World Wide Fund for Nature (1961), Environmental Protection Agency (1970), Greenpeace (1971), United Nations Environment Programme (1972), European Environmental Bureau (1974), Intergovernmental Panel on Climate Change (1988), and European Environment Agency (1993), to name a few.

In addition to the growing public awareness and the emergence of environmental organizations, environmental ideas began to gain political influence in the 1970s. This was reflected in the emergence of green political parties, which draw their ideology from a will to reinforce environmental awareness and protection. Green political parties now exist in many countries around the world. In Germany, for example, the Green Party known as *Bündnis 90/Die Grünen* plays a significant role in German politics. Another example of the European green politics is the French political party *Europe Écologie Les Verts*, whose member Yannick Jadot emphasized climate change as a key issue in his 2022 presidential program, advocating for the use of renewable energy, sustainable modes of transport and agriculture, as well as the protection of biodiversity, reinforcement and reform of environmental education, among other measures (Jadot 2022).

2.3.2 Collective effort

Scientists admit that environmental degradation cannot be combatted without widespread political, corporate and public support, engagement and action. Governments are responsible for developing and promoting environmental policies and

regulations. Companies commit to set emission reduction targets and implement sustainable business practices. Individuals are encouraged to make sustainable lifestyle choices such as reducing personal carbon footprint and waste, conserving natural resources and habitat, and more.

The present situation of environmental mitigation efforts by governments differs significantly across nations. While some countries are leading in implementing sustainable policies and practices, the others lag behind in their efforts. One example of a region that has made significant progress in environmental mitigation is Europe, which has nevertheless acknowledged that much work remains to be done. In 2020, the European Union approved a comprehensive environmental strategy entitled *the European Green Deal*, which aims to make Europe the first *climate-neutral*[6] continent by 2050. To achieve this, the European Green Deal targets to reduce greenhouse gas emissions by 55% and plant 3 billion trees in the EU by 2030. The latest report issued by the Intergovernmental Panel on Climate Change contains the most recent recommendations for governments, including reducing greenhouse gas emissions, fighting against deforestation, restoring natural resources, reforming the food sector, transitioning to renewable energy, promoting energy efficiency, phasing out coal, and encouraging climate action (Harvey 2023).

Businesses, no matter their size, can also have a meaningful impact on environmental mitigation. From a large international corporation reducing its emissions to a small local cafe implementing a zero waste policy, each contribution counts. The Term *corporate environmental responsibility* refers to a company's commitment to sustainable practices in internal processes with regard to waste management, energy efficiency, recycling policy, empolyees' education, and emissions reduction. However, despite these efforts, there is a phenomenon known as *greenwashing*, in which companies falsely claim to be sustainable to increase revenue. For example, the Irish low-cost airline Ryanair has been accused of greenwashing, as its sustainability claims about carbon offsetting were found misleading.[7]

---

[6]The European Green Deal documentation does not offer an explicit definition of the Term *climate-neutral*, leaving the meaning to be inferred from the outlined targets, which emphasize the reduction of greenhouse gas emissions.

[7]https://www.euronews.com/green/2023/01/26/ryanair-low-cost-airline-warned-about-misleading-carbon-offset-claims

Individual actions can add up to governmental and corporate efforts of environmental mitigation. Nowadays, there is a plethora of ways to adopt an eco-conscious lifestyle. While some sustainable tips remain controversial for various reasons, there are universally recognized actions that individuals can take to reduce their environmental impact. For example, transition to plant-based diet can significantly reduce global meat consumption, and as a result, greenhouse gas emissions. Consumerism is another environmental issue that can be addressed through practices such as recycling, buying second-hand items, or participating in "buy nothing" groups[8] where people share things they no longer need for free. Similarly, individuals can reduce plastic waste by avoiding to buy single-use plastic bottles. While many of these environmental actions are voluntary, some are imposed by the government and become obligatory. For example, the recent French law will require French households to compost food waste starting in 2024.[9]

According to a 2020 European Investment Bank climate survey, a significant percentage of Europeans are willing to make sustainable lifestyle changes. 64% are willing to switch to public transport instead of driving, 75% are willing to take less flights, and 78% are willing to reduce the heat consumption (The European Investment Bank 2020).

2.3.3 Environmental communication

In section 2.2, we discussed that environmental science is an academic discipline which studies the natural world from the perspective of the interaction between the nature and humans, environmental issues and ways to address them. In addition to academia, environmental science is also widely promoted by scientists, politicians, journalists, writers and activists to educate the general public about environmental issues and to build environmental awareness. These educators contribute to the collaborative and diversified approach to science communication in order to reach the widest audience possible, as "no single linguistic register is adequate for the totality of discourse needed for addressing the challenge" (Cottey 2022). As for the purposes of environmental communication, Moser (2010: 38) identifies three goals:

[8]For more on *The Buy Nothing Project*, visit: https://buynothingproject.org/.

[9]https://www.service-public.fr/particuliers/actualites/A15940

(i) it serves to educate the general public about the environment; (ii) it aims at promoting public engagement and action; (iii) finally, it encourages deeper changes in social norms, which allow for not only situational environmental actions but also for a fundamental shift in people's minds.

Public engagement and action play a crucial role in mitigating environmental issues. To evoke public response and cultivate the willingess to act on the individual level, it is necessary to educate the general public using a successfull environmental communication strategy. This approach involves creating a meaningful conversation between experts and the general public, assessing the public attitudes toward environmental issues, raising awareness, improving environmental literacy, and more.

However, environmental communication is often regarded as challenging. For example, it can be problematic to communicate complex scientific ideas to the general public in a clear and accessible manner. Furthermore, educators may disagree on the proper methods of communicating environmental science, while politicization can lead to misinformation and ideological bias. Despite these challenges, educators, researchers, and others are working to develop guidelines for more effective environmental communication. These guidelines often center on the language used to communicate environmental science, emphasizing the importance of medium, style, and terminology. While adapting scientific language for the general public, it is critical to preserve factual accuracy. As we are interested in the language of the environment, we will delve into some of these guidelines in greater detail.

Simplified language. In 1992, the American Union of Concerned Scientists published a leaflet entitled *World Scientists' Warning to Humanity*, which aimed to alert people to the fact that human activity was putting the environment at risk (Union of Concerned Scientists 1992). The document can be characterized as a piece of storytelling rather than a solid scientific writing supported by relevant data. Since then, experts have changed their narrative, and they now communicate their alarming findings by backing them up with sophisticated quantitative data and in-depth scientific analysis. As a result, these findings have gained credibility

and validity but have also become more complex for non-experts to understand. Given the lack of scientific knowledge among the general public, environmental science should be communicated in simplified and straightforward messages that are not burdened with complex and unfamiliar terminology. For instance, experts recommend using *human-caused* instead of *anthropogenic* and *time* instead of *temporal* (Hassol 2008). However, simplification should not distort the original message, and the relevance of the expressions used should still be taken into account to promote proper understanding.

Visual language. There are many creative ways to communicate scientific knowledge to the general public and make it more accessible. One effective approach is to use visual presentations of scientific data, which can make complex information easier to understand. An excellent example of this is climate scientist Ed Hawkins' animated graphic called "Spiralling Global Temperatures" (see Figure 2.1), which shows the rise in global temperatures in a minimalistic manner. Beginning near the center of the circle and showing global temperatures in 1850, the line moves in circlular motion, gradually approaching the edges to demonstrate the temperature rise over time. The graphic went viral on social media and was widely featured in mass media as an excellent example of accessible science communication (*Washington Post* 2023). To see more examples of how visualization techniques can make environmental knowledge accessible to a broader audience, one can visit the following sources: *Explaining Climate Change*,[10] *Climate Communication*,[11] *Climate Visuals*,[12] *Skeptical Science*.[13]

Evocative language. Additionally, language can be used to evoke a certain public sentiment or response. While academic discourse is typically neutral and emotionless, it should be "humanized" for a more effective communication (Slovic 2020). To make environmental communication not only simpler, as advised in the first

[10]https://www.explainingclimatechange.com/

[11]https://climatecommunication.org/

[12]https://climatevisuals.org/

[13]https://skepticalscience.com/graphics.php?c=4

[14]https://www.climate-lab-book.ac.uk/2016/spiralling-global-temperatures/

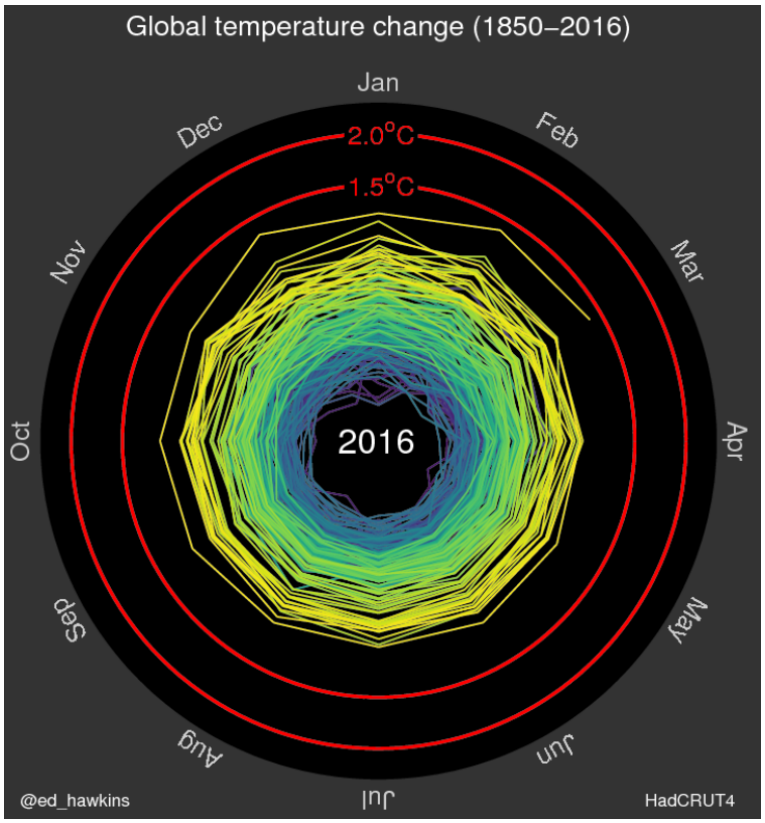

Figure 2.1 – The *Spiralling Global Temperatures* visualization.[14]

point, but also more emotional, one may consider changing the terminology. For example, to emphasize the urgent character of climate change, it is recommended to use expressions such as *climate crisis*, *climate emergency* instead of *climate change* (BBC Science Focus 2020; *The Guardian* 2021). Similarly, Pike, Doppelt and Herr (2010) suggested *climate shock* as a more compelling and evocative alternative. To spark public concern about environmental issues, Krosnick *et al.* (2006) proposed addressing the fact that basic human needs such as food and shelter may eventually be compromised. However, extreme alarmist language should be avoided, as it can "backfire and undermine the intended effects" if not followed by concrete solutions (Feinberg and Willer 2011).

Appealing language. In environmental communication, appealing language is the language which is interesting, hopeful, meaningful and personally relevant. This type of language seeks to emphasize personal impact of environmental issues and avoids presenting them as distant or invisible (Monroe *et al.* 2019). As reported by Lorenzoni *et al.* (2006), if environmental issues are perceived as psychologically distant, they are not viewed as a primary concern. Rather than just presenting

scientific facts, appealing language uses storytelling to engage the audience: "most people understand the world through anecdotes and stories, rather than statistics and graphs" (Corner, Shaw and Clarke 2018). Furthermore, appealing language should be hopeful and propose solutions, along with explaining what actions individuals can take to mitigate environmental issues and the benefits of doing so.

Language of certainty. In 2006, the discourse on climate change was described as "confusing, contradictory and chaotic" (Ereaut and Segnit 2006). Despite efforts to promote settled environmental science, uncertainty remains a significant part of the modern environmental discourse. To avoid further compromising the scientific consensus and spreading skepticism, it is recommended to avoid confusing language such as *debate*, *risk*, *uncertainty* (Hassol 2008). Uncertainty negatively affects public willingness to engage in environmental discussion and take action (Patt 2007). According to D. Wood and Vedlitz (2007), the public tends to be more involved with environmental topic when they know that scientists agree and have a thorough understanding of environmental issues.

Metaphoric language. Metaphors[15] play an important role in popularizing environmental science, as they make complex scientific ideas more accessible. For example, Hassol (2008) proposes using a "pot of water" metaphor to explain how climate predictions work. People often wonder how it is possible to predict the change in climate in the next decades while even short-term weather forecasts fail to provide accurate data. The metaphor suggests to look at climate predictions through analogy with the process of boiling a pot of water: it is impossible to predict exactly where and when the first bubble will form, but you can predict that the water will boil in about 10 minutes. As can be seen, metaphors make dry scientific facts more catchy. Furthermore, metaphors are effective in drawing public attention. As asserted in Ungar (2000), the ozone hole attracted a wider public attention in the late twentieth century because it was covered in press with the use of metaphors, while it was not the case for climate change. It should be noted, however, that inappropriate metaphors can have the opposite effect such as

[15]Metaphor is a rhetorical figure which consists in understanding one concept in terms of another concept (Lakoff and Johnson 2003).

causing confusion and misunderstanding.

Targeted language. Although building environmental awareness is a global undertaking, communication practices should be adjusted to reach different social groups based on their geographical area, environmental views, cultural and social values, and level of scientific literacy. Targetting different audiences with adjusted language and relevant examples can improve the effectiveness of environmental messages. For example, educating people about environmental issues which are locally relevant to a given group might resonate better with them.

Despite the huge number of information channels aimed at raising environmental awareness, communication efforts still seem to be insufficient (Cottey 2022). Studying public sentiment about the environment can help construct a better communication strategy, particularly given the alarming level of uncertainty and mistrust among the general public. Effective environmental communication is not a universal remedy, as noted in Nerlich, Koteyko and Brown (2010). A combined effort is necessary, of which communication is only a part, to take environmental issues under control.

The Intergovernmental Panel on Climate Change (IPCC) is a large international body responsible for synthesizing and disseminating the latest scientific knowledge on climate change. Its goal is "to provide governments at all levels with scientific information that they can use to develop climate policies."[16] The IPCC is a powerful voice of environmental communication. It produces various document genres, including summaries for policymakers, whose target audience is political leaders. However, it is often criticized for the language or communication techniques (Bowman *et al.* 2009; Benestad 2023).

For example, the organization uses a specific terminology in its reports to indicate the level of confidence in the scientific statements. The Term *likely* corresponds to 66-100% of assessed likelihood:

> Global warming is *likely* to reach 1.5°C between 2030 and 2052 if it continues to increase at the current rate. (IPCC 2018)

[16]https://www.ipcc.ch/about/

However, this approach contributes to even greater confusion among the readers, as the general public has their own interpretations of the likelihood markers (Budescu *et al.* 2014).

2.3.4 Uncertainty and mistrust

As stated previously, the dialogue between scientists, educators, and the general public has an important strategic role in addressing environmental issues. Public trust in environmental science is a key component of this dialogue, as it has a strong correlation with public willingness to participate in environmental mitigation. The more people believe in scientists' warnings, the more likely they are to take necessary actions. Furthermore, trust is also a major factor of public sense-making when there is a lack of solid environmental knowledge.

However, one of the main challenges to public trust has been the idea that environmental science is not settled, and there is uncertainty and disagreement among scientists' claims. As Krosnick *et al.* (2006) argued, people who are aware of scientific uncertainty around global warming tend to be more skeptical about the topic. Uncertainty, contradictory findings, disagreement, and debate have always been an integral part of the scientific process, and environmental science is no exception. While there may be contradictions in scientific claims about environmental issues, it must be noted that there is a strong evidence-based scientific consensus. 97% of environmental scientists agree that human activity is causing global warming (Cook 2016). In 2007, the Intergovernmental Panel on Climate Change reported that "warming of the climate system is unequivocal" (IPCC 2007). In 2021, IPCC reaffirmed its claims by stating that "it is unequivocal that human influence has warmed the atmosphere, ocean and land" (IPCC 2021). The scientific community continues to admit the existence of environmental dangers and insists on their imminent character:

> *The world is in a climate emergency – "a code red for humanity" according to the UN Secretary-General* (UNEP 2021).
>
> *The science is clear. Climate change is real. Climate change is happening now* (WWF 2023).

> *There is unequivocal evidence that Earth is warming at an unprecedented rate. Human activity is the principal cause* (NASA 2023a).
>
> *Climate change affects us all and is accelerating* (EEA 2023).
>
> *Climate change is a very serious threat, and its consequences impact many different aspects of our lives* (European Commission 2023b).

Despite this scientific consensus, public opinion on the importance and existence of environmental issues remains polarized. Generally speaking, there are two groups: those who believe in the existence of environmental issues and those who doubt or do not believe at all. The latter group is usually referred to as skeptics or deniers.[17] The level of public skepticism, and even denial, remains a matter of concern. Clearly, the proportion of skeptics and deniers varies with different communities and fluctuates over time. According to one of the latest surveys conducted by the French electric company Électricité de France[18] (EDF), 66% of the global population considers the climate change a major issue (EDF 2022). Although the level of concern is high (69%), it tends to decrease (as compared to 72% in 2021). On the contrary, the level of skepticism has increased from 31% in 2019 to 37% in 2022.

Although scientists are generally regarded as trustworthy, there are several variables that can influence an individual's sense-making of scientific claims. These include prior knowledge, exposure to mass media (which can provide bias or false information), political views, personal experiences, and social environment. For example, research by Krosnick *et al.* (2006) found that people's belief in global warming was strongly influenced by their personal experiences, such as witnessing the rise in local temperatures. On the contrary, the absence of noticeable changes in weather patterns can be interpreted as a proof that climate change is not happening. Political views can also influence the level of public trust. In the United

[17] In the context of environmental debate, the definition of *skeptics* and *deniers* vary with source. In general, skeptics express uncertainty about the existence of environmental issues, in particular climate change, and usually do not admit the fact that it is caused by human acitivity. Deniers, on the other hand, deny the existence of human-caused environmental issues and their impacts on the society. The denial movement draws on the belief that scientific claims regarding environmental issues are not reliable and are uncertain.

[18] **En.** Elictricity of France

States, environmental debate is often politicized. According to Hmielowski *et al.* (2014), conservative American media tend to undermine public trust in scientists and their findings, while exposure to non-conservative American media contributes to the increase of trust and certainty in scientific claims.

While skepticism and denial can stem from a range of factors, the blame is often on scientists who "have not been effective communicators", as noted by Hassol (2008). Given the urgent need for collective effort in environmental mitigation, bulding public trust should be a priority for educators.

2.3.5 Environmental literacy

As members of the modern society, we encounter scientific concepts on a daily basis, including those related to the environment. The scientific literacy of the general public has long been studied and researchers are specifically interested in measuring the level of public literacy with regards to the environmental topic, as it has direct implications on the ability to form informed opinion and to take actions.

The Declaration on the human environment adopted by the first United Nations Conference on the Human Environment in Stockholm, Sweden in 1972 is an important document for environmental literacy. One of its principles reads as follows:

> Education in environmental matters, for the younger generation as well as adults, giving due consideration to the underprivileged, is essential in order to broaden the basis for an enlightened opinion and responsible conduct by individuals, enterprises and communities in protecting and improving the environment in its full human dimension. It is also essential that mass media of communications avoid contributing to the deterioration of the environment, but, on the contrary, disseminate information of an educational nature on the need to protect and improve the environment in order to enable man to develop in every respect.

In this passage, the authors emphasize the importance of environmental education as the foundation for informed environmental opinions and delegate the responsibility of disseminating environmental knowledge to the mass media.

What is environmental literacy in the modern world? The North American Association for Environmental Education defines environmental literacy as follows (Hollweg *et al.* 2011):

> Environmental literacy is knowledge of environmental concepts and issues; the attitudinal dispositions, motivation, cognitive abilities, and skills, and the confidence and appropriate behaviors to apply such knowledge in order to make effective decisions in a range of environmental contexts. Individuals demonstrating degrees of environmental literacy are willing to act on goals that improve the well-being of other individuals, societies, and the global environment, and are able to participate in civic life.

The notion can be further narrowed down to address specific environmental issues, like climate change, resulting in the idea of climate literacy. U.S Climate Change Science Program (2009) defines a climate-literate person as someone who knows how the Earth's climate system functions, is capable of distinguishing credible scientific information from unreliable claims, can communicate properly about climate change, and can make informed contributions to the mitigation of climate change.

Environmental science is a complex field, covering many different disciplines, making it difficult for the general public to navigate and understand the scientific basis. While scientists develop their understanding of environmental issues through rigorous scientific methods, the general public's understanding of environmental science often relies on personal experience which makes them more susceptible to misunderstanding (Weber and Stern 2011). This is further complicated by the fact that politicians and journalists, who are the primary sources of information for the general public, are not always reliable and may provide misinformation. Being aware about misinformation techniques reduces the susceptibility to such information which is another point of improvement in how people are educated about the environment (Schmid-Petri and Bürger 2022).

Insufficient public knowledge about the environment can have complex social and political consequences. For instance, a lack of knowledge and subsequent skepticism towards new technologies could delay their adaptation. To avoid potential

ignorance and prejudice against emerging technologies, it would be beneficial to involve the general public's opinions in the early stages of development through a process known as "upstream public engagement" (Wibeck *et al.* 2017). The notion of upstream public engagement primarily refers to emerging technologies under development. Nevertheless, involving the general public in scientific debate should also be applied to established scientific concepts and technologies which may have been already distorted in the public understanding, in order to provide clarity.

To study what people know about the environment and environmental issues, researchers use methods like surveys and interviews. Current research analyzes the general public's perceptions of environmental changes in order to adapt communication campaigns and mitigation policies according to existing knowledge. Previous studies have identified several pain points in public understanding of the environment, including oversimplification of complex issues (Coyle 2005), misunderstanding of the basic physical processes (Sterman and Sweeney 2007), and difficulty engaging in meaningful conversations about the topic (Henry 2000; Asiyanbi 2015). Additionally, people struggle to establish causal relationship, to distinguish between movie fiction and science facts in movies about environmental disasters (Lowe *et al.* 2006), and are uncertain about how to take actions (Glasgow *et al.* 2018).

Educators report on the following common misconceptions related to environmental science, among many others:

- There is a belief held by some that climate change is entirely natural and that the Earth's climate has always been changing. However, the scientific communitry agrees that human activity contributes to the increase in greenhouse gas concentrations, which has a significant impact on the climate (NASA 2023b).
- It is a common misconception that renewable energy sources are not reliable and are just as harmful to the environment as traditional sources of energy. In fact, renewable energy sources emit from 14 to 134 lower amounts of greenhouse gases compared to fossil fuels (WWF 2014).
- Some mistakenly believe that carbon dioxide is the only greenhouse gas, likely due to the special attention to it from scientists and the media. While

> carbon dioxide is indeed a potent greenhouse gas, there are other greenhouse gases which also contribute to the excess greenhouse gas emissions, such as methane, nitrous oxide, and fluorinated gases (NewScientist 2007).

As cited by McCaffrey and S. M. Buhr (2008), and originally proposed by Gowda, Fox and Magelky (1997), such misconceptions might arise due to the following reasons: (i) limited acces to reliable climate change information; (ii) incomplete and distorted news media coverage; (iii) public interpretation of climate science based on intuitive judgements; and (iv) uninformed environmental activism which stems from inability to determine cause-and-effect relationship between environmental phenomena.

In summary, *environmental literacy* among general public enables people to make informed decisions, evaluate the reliability of environmental news, and contribute meaningfully to the public debate. Additionally, environmental literacy influences the public ability to understand environmental Terms and use them correctly.

2.4 Chapter synthesis

In this chapter, we have provided an overview of the environment as a topic that has gained international prominence over the last 50 years. Although its recognition as a top global issue was not straightforward, nowadays the importance of environmental debate is unquestionable. Terms such as *the environment* and *climate change*, as well as many others, have become household names. The notion of the environment refers to the Earth's natural world that is being destroyed, polluted, and harmed by human activity, and needs protection and conservation.

Although the environmental domain is rooted in scientific evidence, it has strong social implications and now extends beyond academic concerns. Society as a whole, and every individual, who are affected by current and emerging environmental issues, play a critical role in preventing them by showing empathy towards the topic and willingness to act. There are still many challenges to address, including public mistrust, lack of collective effort, ineffective environmental communication, misinformation campaigns, political bias, and a lack of scientific

knowledge among the general public.

In the following chapter, we will shift from the historical and social perspective on the environmental domain to a linguistic analysis.

Chapter 3

Terminology of the environment

Summary

3.1 Notion of terminological unit

One of the most important aspects of our research is the study of terminological units. This section presents fundamental terminological notions, including the opposition between specialized and general language communication and, by extension, between specialized lexical units (Terms) and general language lexical units. It also covers different approaches to terminology, including a description of the theoretical approach adopted for this research. Finally, we propose a new terminology for the classification of Terms with respect to the process of domestication.

3.1.1 Specialized domain and specialized communication

In this section, the discussion is built mainly around the opposition and interconnection of two vast notions, *specialized communication* and *general language communication*, or non-specialized communication. This opposition draws on several principal distinctions between two types of communication (this list is not exhaustive):

- Specialized communication and general language communication are characterized by the usage of distinct linguistic tools: Terms, as specialized lexical units conveying domain-specific information, and general language lexical units which do not normally convey any domain-specific information.
- As opposed to general language communication, specialized communication has a very specific purpose, that is, communicating and transmitting specialized knowledge.
- Specialized communication is characterized by specific thematic boundaries of a given specialized domain, whereas general language communication is not subject to such restrictions.
- Specialized communication requires training in a specific subject and mastering of its terminology. As a consequence, the main actors of specialized communication are domain experts. In contrast, general language communication does not imply any specific training.

- Specialized language is grammatically restricted as compared to the general language as it uses only a part of the general language grammatical tools which concern morphology, syntax, semantics, etc. (Kittredge 1982).[1]

This list contains a number of key notions which require definitions. Let us begin with the notion of specialized domain. *Specialized domain*[2] is an area of activity or knowledge that exists within a limited thematic context. Typically associated with scientific disciplines like linguistics, mathematics, or physics, this notion also encompasses domains unrelated to science, such as leisure or professional activities like gardening, cooking, TV show production, football, and others. Each specialized domain has its own system of lexical units that communicate knowledge specific to that domain. Such lexical units are called Terms.

3.1.2 Term, with a capital *T*

The notion of *Term* is a keystone of terminological studies and has thus received significant attention from terminologists, who never cease searching for a "perfect" definition of Term. If two terminologists were given a specialized text and asked to produce a list of Terms, they would generate two different lists of potential Terms, unless they were given specific instructions regarding the task's objectives. However, even with clear instructions, the resulting lists would still differ for several reasons. Firstly, there is no single, universally accepted definition of a Term among terminologists. Secondly, different terminological schools and approaches propose different methodologies for treating Terms – see, for example: Rey (1979), Rondeau (1984), Condamines (1995), Temmerman (2000), Kageura (2002), Humbley (2018), Cabré and Azevedo (2019), Collet (2021). Finally, the notion of Term can be defined and fine-tuned according to the objective of a given research.

Despite numerous existing definitions of Term, we adopt the approach that considers Terms as lexical units,[3] as exemplified by the contributions in the special

[1] In Kittredge (1982), the author uses the notion of *sublanguage* which is synonymous to our conceptualization of specialized language.

[2] Further in the text, we use the notions of specialized domain and domain interchangeably.

[3] The notion of lexical unit is introduced in section 1.3.

issue of the *Terminology* journal: Faber and L'Homme (2014). We propose the following definition of Term:

> A *Term* is a specialized lexical unit which refers to a domain-specific concept and serves for transmitting domain-specific information.

In literature, the notion of Term is usually illustrated with science-related lexical units. However, let us consider an example from poker domain. Poker is a specialized domain within the broader area of card games. To convey information related to poker, we use specific poker Terms (highlighted in bold):

(1) You need to always know what **pot odds** you are getting if you want to make good **calls** consistently.
[Web, https://upswingpoker.com/pot-odds-step-by-step/, 23/06/2023]

These are poker Terms which contain specialized knowledge specific to poker. In poker, the noun *call* denotes a bet made to continue playing, and the phraseme *pot odds* denotes the ratio between the total amount of bets made by all players and the amount a player must bet at a given moment. Without mastering these and other poker Terms, it is impossible to play the game or communicate effectively about it. Mastering terminology requires the ability to use Terms correctly and appropriately within their semantic and grammatical contexts. For example, to become proficient in poket terminology, one can gain first-hand experience in playing the game or develop an interest in it (watching games, studying the rules, etc.). On the other hand, mastering terminology of a scientific field requires education in a respective scientific discipline.

It is also important to note that different domains can share Terms (Delavigne 2002). For instance, the Term *microscopic* is used in chemistry, physics, biology, medicine, and other fields. Throughout this text, we will indicate the domains of each Term within its given context. However, it is worth keeping in mind that a single Term may be shared by multiple domains within a broader context.

In previous research, Terms are sometimes described in opposition with *words*, as discussed in Adelstein and Cabré (2002), Halskov (2005), Humbley (2009), and S. A. Khan (2016). Words are usually associated with the units of lexicographic

and lexicological description. On the other hand, Terms are associated with terminographer's work. We do not endorse the use of the notion of word as it is ambiguous and thus unsuitable for technical writing. Furthermore, we do not consider *Terms* and *words*, or more accurately general lexical units, as mutually exclusive categories. In fact, the opposition is rather between specialized lexical units (Terms) and general lexical units.

Important writing convention. To ensure greater clarity in our linguistic analyses, we introduce an important writing convention: we use the capital *T* to denote specialized lexical units that we consider as "true" specialized lexical units. In the following examples, we refer to specific terminological senses: 'chemical element' in (2) and 'minute marine orgamisms' in (3).

(2) We finished working on the **Term** CARBON I in the English Lexical Network.

(3) The **Term** *phytoplankton* belongs to the terminology of biology.

In contrast, we do not capitalize the notion of term when we do not refer to any specific terminological unit, but we are still discussing terminology in general. In a sense, we are referring to a terminological "word". Consider the following example:

(4) The **term** *carbon* is often used in general language.

What do we actually mean by *term carbon*? Are we referring to the chemical element carbon, the umbrella Term for greenhouse gases, or something that the general public considers a Term but is not actually one? The introduced writing convention allows us to be precise in detailed linguistic analyses. However, it also allows us to stay at the level of abstract generalizations without going into detail when needed.

3.1.3 Approaches to terminology

We differentiate between *terminology*$_{\mathbf{1}}$ as an organized and coherent system of all Terms of a given domain and *terminology*$_{\mathbf{2}}$, which refers to the study of terminologies$_{\mathbf{1}}$,

with different schools of thought promoting distinct approaches and solutions for terminology management.

The General Theory of Terminology, authored by the Austrian terminologist Eugen Wüster (Wüster 1991), is one of the most widely recognized terminological theories and is known as *traditional* or *classic terminology*. The General Theory of Terminology aimed to stardardize and disambiguate the terminology to facilitate technical communication. It adopted onomasiological approach which manifests itself in the idea that concepts should serve as the starting point for terminological description. *Concept* is viewed as an extralinguistic mental construct which is designated by a linguistic element *Term*. The onomasiological approach prioritizes the conceptual system as the foundation of specialized domains but does not fully consider the linguistic aspect of specialized discourse and the complexity of Terms.

However, advances in terminology$_{\mathbf{2}}$ have highlighted the need for a semasiological approach to terminology$_{\mathbf{1}}$ treatment, by placing Terms at the center of terminological study. With the appearance of corpus linguistics, it has become possible to track Terms in rich text corpora, paying special attention to their lexical environment. This information has made it possible to construct a more complete linguistic profile of any Term, which traditional terminology overlooked.

Access to a large amount of data has provided a fertile ground for the lexico-semantic approach to terminology, which focuses on the semantics and linguistic behavior of Terms in texts. Various semantics-oriented theories, such as Cognitive Semantics (Talmy 2000), Frame Semantics (Fillmore 2006), and Explanatory Combinatorial Lexicology (Mel'čuk 1995), are now applied in terminologists' work. This approach has proved successful in analyzing specialized languages, as seen in the following works: Jousse and Bouveret (2003), Dancette and L'Homme (2004), L'Homme (2018), Bonadonna (2020).

Lexical-semantic approach has been also applied in the context of environmental terminology. For example, the environmental dictionary *DiCoEnviro*,[4] developed at the Observatoire de linguistique Sens-Text,[5] is a multilingual[6] lexico-

[4]http://olst.ling.umontreal.ca/dicoenviro

[5]**En.** Meaning-text Linguistics Observatory

[6]DicoEnviro is available in languages such as French, English, Spanish, Italian, etc.

graphic source that adheres to the principles of Explanatory and Combinatorial Lexicology (L'Homme 2015; L'Homme 2016). Specifically, it treats environmental Terms as lexical units, with an emphasis on their specialized meaning and lexical relations with other Terms.

Proponents of the lexical-semantic approach to terminology have pointed out the disadvantages of traditional practices and proposed solutions to improve terminological description. Specifically, they agree on three points of criticism explained below (A-C).

A) Monosemy *vs.* polysemy. The traditional approach to terminological description emphasizes clarity and explicitness by focusing on monosemy and asserts that a Term can only be associated with one concept. However, in reality, one Term (or rather one linguistic designation) can refer to several concepts. Therefore, polysemy, as well as synonymy, should not be disregarded, but rather considered essential aspects of terminological description.

Within one terminological polysemous vocable, polysemy can manifest itself in the following ways:

- **Polysemy within one specialized domain:** In medical discourse, *asymptomatic* can take different arguments depending on the context, which makes it necessary to describe two separate senses:[7] ASYMPTOMATIC **1**, which denotes an illness that has no signs (*asymptomatic flu*), and the metonymic Term ASYMPTOMATIC **2**, which denotes a person who suffers from such an illness (*asymptomatic patient*).
- **Polysemy that covers several specialized domains:** In medicine, the Term *diagnosis* is used to denote a medical examination of an ill person in order to identify the nature of the illness (*diagnosis of a disorder*). In biology, *diagnosis* is used to denote a taxonomic description of an organism (*diagnosis of the species*). Therefore, the vocable DIAGNOSIS contains both the medicine- and biology-related senses.

[7]The numbering of lexical units is symbolic and does not account for complete polysemy of the respective vocable.

- **Polysemy that covers both specialized domain and general language:** The physical Term *vacuum* had been adopted by the general language where it has acquired new senses. For example, one of the general language senses of *vacuum* denotes a metaphorical situation in which the absence of someone or something creates an abstract emptiness (*political vacuum, ideological vacuum*).

The notion of polysemy goes in hand with the notion of copolysemy. As asserted in Polguère (2018), *copolysemy* refers to the semantic relation between two lexical units within one polysemous vocable. Lexicographic analysis allows us to examine copolysemy relations between lexical units within one vocable and to identify the type of such relations: *Metonymy*, *Metaphor*, *Extension*, *Generalization*, *Specialization*, etc. For example, the physics Term *vacuum* and its general language counterpart are connected with copolysemy link which is based on metaphor.

B) Conceptual hierarchy *vs.* non-hierarchical system. Traditional terminology recognizes hierarchical relations between concepts, such as generic-specific (*measles* is a *disease*) and part-whole (*motherboard* is a part of a *computer*), which allows for the building of a hierarchical conceptual structure. However, relations between Terms are more complex, and non-hierarchical relations should be also accounted for. *Explanatory and Combinatorial Lexicology* proposes a formal tool called *lexical functions* (Mel'čuk and Polguère 2021) to model a wide range of lexical relations in natural languages:

- Simple lexical functions:
  - `Syn`(*carbon dioxide*) = CO_2 (exact synonym)
  - `Gener`(*carbon dioxide*) = *gas* (generic name)
  - $\texttt{S}_{\texttt{loc}}$(*carbon dioxide*) = *atmosphere* (typical location)
- Complex lexical functions:
  - `CausPredMinus`(*carbon dioxide*) = *cut, reduce* ('to cause to diminish in quantity')
  - $\texttt{LiquFunc}_0$(*carbon dioxide*) = *sequester* ('to make stop functioning')

C) Non-nominal Terms. The traditional approach primarily focuses on nouns and tends to ignore other parts of speech. This is because, in traditional terminology, concepts are primarily associated with entities which are linguistically expressed by nouns or noun phrases. However, specialized knowledge also include other semantic classes, such as events, properties, relations, and states, expressed by verbs, adjectives, and adverbs. L'Homme (2003) argues that analyzing specialized verbs can be a good starting point for discovering the lexical structure of a subject field, owing to their predicative nature and ability to take arguments. For example, the environmental verb *recycle* takes three arguments: 'X recycles Y into Z' (*the* $machine_X$ *recycles green* $waste_Y$ *into* $compost_Z$).

3.2 Migration of Terms into general language

3.2.1 Domestication *vs.* determinologization

Although specialized language is a category on its own, it does not exist in isolation. Specialized language and general language interact on several levels. First, any piece of specialized discourse (either in written or oral form) is composed of both specialized and general lexical units which coexist harmoniously in a single linguistic environment. Second, specialized language and general language sometimes exchange their lexical stock.

Such lexical exchange can be attributed to various factors and entail different consequences depending on its orientation. On the one hand, Terms can migrate from specialized language into general language. On the other hand, general lexical units can also move into specialized language, a process known as *terminologization* (Dury 2008). As asserted in Meyer and Mackintosh (2000), IT terminology often reuses general lexical units such as *mouse*, *mailbox*, *menu*, *wallpaper*. In Dury (2008), the author traces the appearence of general-language *carbon neutral* in specialized discourse. Another process, referred to as *transterminologization*, involves transferring Terms between different specialized domains. For example, Gorbunova, Archakov and Akhtayeva (2020) studied the terminological exchange between the domains of programming and stock trading.

We concentrate on the first kind of lexical exchange, namely the migration of

Terms into general language. In linguistic studies on terminology, this process is most commonly known as *determinologization* (Meyer, Varantola and Mackintosh 1998; Meyer and Mackintosh 2000; Nová 2018), but it is also discussed in related but more specific contexts, such as popularization (Galisson 1978; Delavigne 2003; Vargas 2009; Botta 2013; Moirand 2017; Ledouble 2020), despecialization (Condamines and Picton 2014b), and neology (Costa *et al.* 2022; Humbert-Droz 2023), among others.

According to the theoretical framework described in Meyer and Mackintosh (2000), there are two main types of determinologization, as illustrated by examples provided in the article:

- In the first type, the domain-specific sense of the Terms is (mostly) preserved. For instance, medical Terms *HIV*, *bulimia*, *MRI* refer to the same concepts for both experts and non-experts, although non-experts may have shallower or connotated perceptions compared to experts.
- In the second type, the domain-specific sense of the Term is significantly altered. The semantic change is so substantial that the newly developed general language sense no longer refers to the same concept. For example, in the context of waste-management, *recycle* refers to 'processing utilized objects to create new ones', whereas in general language *recycle* can apply to abstract concepts like ideas, screenplays, and films, as seen in *recycling of musicals*.

Nová (2018: 391) proposed a more nuanced typology, introducing an intermediate case between sense presevation (referred to as *type A*) and sense alteration (*type C*) – the blurring of the sense (*type B*). This type suggests a significant change in the Term's meaning but such a change is not sufficient to define a completely new sense in the general language. As an example, the author discusses the psychiatric Term *depression* that refers to a mental condition in specialized discourse and is commonly used to describe feelings of sadness in general language discourse.

In both typologies discussed (Meyer and Mackintosh 2000; Nová 2018), as well as in other works on determinologization (Halskov 2005; Gorokhova 2021),

the most extreme cases of determinologization, where new senses emerge in the general language, are often illustrated by figurative senses based on metaphor or metonymy. Consider the following examples of such metaphoric senses and their scientific counterparts:

- IT Term *reboot* 'restart a program' and its general language equivalent *reboot* 'start something again' (Halskov 2005);
- medical Term *allergy* 'disorder of the immune system' and its general language equivalent *allergy* 'strong dislike of something' (Nová 2018);
- engineering Term *drill*$_{(N)}$ 'industrial machine for boring' and its general language equivalent *drill*$_{(N)}$ 'repeated exercise' (Gorokhova 2021).

In the context of determinologization, we approach Term migration from a slightly different perspective. Specifically, we focus on the *domestication* of Terms in the general language, considering changes that occur in semantics, combinatorics and pragmatics due to uninformed usage of specialized vocabulary by non-experts. We will discuss the types of domesticated Terms in more detail in 3.3.

3.2.2 Ordinary discourse

One of the most important channels for transmitting specialized knowledge from experts to the general public is mass media. Despite this intermediary role, mass media discourse is often considered as non-specialized within the context of determinologization. While specialized language corpora mainly consist of scientific textbooks, papers, articles, PhD theses, and similar sources, general language corpora are often built with mass media texts such as newspaper and magazine articles, as can be seen in Dury (2008), Botta (2013), Condamines and Picton (2014a), and Humbert-Droz and Picton (2022). As pointed out by Botta (2013), journalists are not experts and may interpret Terms differently as compared to scientists. Furthermore, Dury (2008) introduces the notions of *word* and *general lexical unit* in the context of newspaper language. Similarly, Condamines and Picton (2014a) argues that general press (French magazines like *Le Monde* and *Le Figaro*), as opposed to popular science press, serves as a valuable source for studying Terms in general language discourse. However, popular science texts are also studied in

opposition to specialized discourse (Nicolae and Delavigne 2009). Humbert-Droz and Picton (2022) used popular science texts as a corpus of intermediate expertise.

Additionally, research on Term migration focuses on more informal registers than the general press, such as general public communication. An accessible way to study this type of communication is by analyzing texts found on online platforms built for social interactions, such as *Twitter*, *Reddit*, *Facebook*, *Instagram*, to mention a few. Users of these social networks generate a vast amount of textual data, such as posts and comment threads, every second, making social media data a rich source for studying general public communication.

In social research, the notion of *general public* (also referred to as *citizens* or *lay people*) refers to individuals with no scientific expertise, as opposed to scientists who are considered knowledge bearers. It is important to note that due to the heterogeneous demographics of online users, not all of them can be categorized as lay people with the same level of knowledge. However, we assume that the majority of people lack expertise in environmental science. On the other hand, we worked with Reddit data posted in thematic environmental communities that might attract more experts and semi-experts than non-specialized communities.

As we are interested in domestication use of Terms, we primarily aim to study online posts of non-expert users on Twitter and Reddit. Nevertheless, the presence of potentially expert-written texts in our data is not a significant concern. In certain communicative contexts, experts often minimize the use of Terms or replace them with colloquial alternatives, which is an interesting linguistic phenomenon on its own. For example, León-Araúz (2017) discusses how doctors might employ different language in a doctor-patient situation as compared to a medical conference. Similarly, we believe that the informal nature of social networks indirectly influences the discursive behaviour of users, including experts, who might adapt to the non-specialized environment when they use such platforms. Studies by Gero *et al.* (2021) and Giusta, Jaworska and Greetham (2021) show that experts tend to use subjective and informal language on Twitter.

Given these points, we use the notion of *general public* to refer mainly to non-expert citizens in a broad context, and predominantly non-expert online users in context of social media. However, we acknowledge the potential heterogeneity

of discourse on social networks and in our corpus in terms of expertise level.

To conclude, previous studies on determinologization use the notion of general language discourse to refer to either general press or general public communication, or both. We adopt the latter approach and use *general language discourse* as an umbrella expression to encompass both general press and general public communication. However, our research primarily focuses on studying environmental terms in texts found on social networks. Therefore, we use the notion of *ordinary discourse* that refers to linguistic productions of the general public on social networks, such as *Twitter* and *Reddit*. A more detailed discussion of social networks and ordinary discourse will be provided in Chapter 4.

3.3 Types of Terms

Domestication process will be further discussed from the perspective of three notions:

- *Full Terms* are specialized lexical units reserved for specialized communication between experts.
- *Runaway Terms* are full Terms which moved into general language but preserved their domain-specific sense.
- *Quasi-Terms* are non-Terms which originate from incorrect/loose interpretation of full Terms by non-experts. Such non-Terms can no longer be considered Terms due to significant semantic and grammatical mutations.

We differentiate between these three types of (quasi-) terminological units[8] based on their semantic properties and the kind of discourse they are found in. In the English Lexical Network, there are three linguistic usage notes to label each type of terminological unit: "*spec*" for full Terms, "(*spec*)" for runaway Terms and "*quasi-spec*" for quasi-Terms. The following sections give a detailed overview of each type accompanied with illustrative examples.

[8]Cf. Mikhel (2022) to see how these notions are applied to the chemistry terminology.

3.3.1 Full Terms: **spec**

Characterization. A full Term is a specialized lexical unit which has a very specific communicative function: it is used by domain experts to communicate domain-specific knowledge. It is not commonly used in general language for several reasons. First, the general public is often unaware of such Terms because they refer to domain-specific knowledge that is unknown to laypeople. Second, even if people are aware of these Terms, they may lack the specialized knowledge necessary to use them correctly. Finally, there may be little or no need to use such Terms in everyday language. Consequently, full Terms are typically used in specialized discourse by experts in the corresponding field(s).

Illustration. The Term *phytoplankton* belongs to the field of biology and refers to small marine organisms. Mastery of this Term requires scientific expertise in biology, as well as knowledge of related Terms such as *algae*, *cyanobacteria*, *diatoms*, *dinoflagellates*, which comprise phytoplankton. Without this specialized knowledge, it is difficult to use the Term correctly, and there may be little need to do so outside of scientific communication.

Lexicographic treatment. Full Terms are usually found in specialized dictionaries, but they may also be included in general dictionaries. In such cases, authors may either avoid using other full Terms to define them and provide simplified definitions, or they may use full Terms to provide more precise definitions, similar to those found in specialized dictionaries. To illustrate, let us compare the definitions of *phytoplankton* taken from a specialized dictionary (5) and two general dictionaries (6), (7):

(5) The photosynthesizing plankton, consisting chiefly of microscopic green algae, diatoms, dinoflagellates, and cyanobacteria. [...] [*Oxford Dictionary of Biology*, 2015]

(6) Plankton that consists of minute plants and other photosynthetic organisms, including cyanobacteria, diatoms, and dinoflagellates. [*American Heritage Dictionary*, 2023, online]

(7) Planktonic plant life. [*Merriam-Webster Dictionary*, 2023, online]

Although *Oxford Dictionary of Biology* is a specialized dictionary and *American Heritage Dictionary* is a general dictionary, the definitions (5) and (6) are almost identical. Both dictionaries define *phytoplankton* with biological Terms: *photosynthesizing*, *photosynthetic*, *algae*, *diatoms*, *dinoflagellates* and *cyanobacteria*. On the contrary, example (7) is a derogatory definition which does not include any of the above-mentioned Terms.

In addition, general dictionaries typically indicate domain-specific notes for specialized lexical units. For example, the *Oxford Lexico Dictionary* marks *phytoplankton* with the usage note **biology** to signal that it belongs to the domain of biology. However, neither the *American Heritage Dictionary* nor the *Merriam-Webster Dictionary* provide any such label for this Term. Within the English Lexical Network, full Terms are designated with the usage note **spec**.

As asserted in (Polguère 2023: 18), it is unnecessary to have a domain label, such as **biology**, in a Lexical System due to several reasons. First, the usage note **spec** already provides information about the terminological status of a lexical unit. Second, the related lexical units that surround a particular specialized lexical unit in a lexical network indicate the domain to which they belong. Therefore, biology-related lexical unit is embedded within a lexical network of biology Terms.

3.3.2 Runaway Terms: **(spec)**

Characterization. As opposed to full Terms, runaway Terms are specialized lexical units which are used by both domain experts and the general public. The metaphoric expression *runaway Term* implies that these Terms have escaped from specialized discourse into general discourse and can be found in both. In terminological research, such Terms are also known as *sub-technical terms* (Marín and Rea 2014) and *common terms* (Rigouts Terryn, Hoste and Lefever 2018).

On semantic level, the domain-specific sense of runaway Terms remains mostly intact in general language contexts. This vision aligns with the concept of *retention of domain sense* in Meyer and Mackintosh (2000) and *type A* in Nová (2018). When a Term preserves its terminological sense in general language, there can be differences in how well the general public understands it, with some having a superficial perception of the Term, while others might be able to use it more

knowledgeably. However, we do not observe significant shifts in the meaning. Therefore, we do not consider it necessary to register the uses with minor semantic deviations as new general language senses.

Illustration. To delve into runaway Terms in more detail, let us first reflect on the recent outbreak of COVID-19 which introduced or reinforced the use of numerous medical Terms in everyday language, such as *asymptomatic*, *antibody*, *coronavirus*, *herd immunity*, *pandemic*, *incubation period*, *quarantine*, *vaccine*, etc. The Term *vaccine*, which was already quite commonly used in general language before the recent coronavirus pandemic, received special attention due to massive vaccination campaigns organized in many countries. In medical discourse, *vaccine* refers to a medical product that is introduced into the human body to stimulate the immune system and to make the body resistant to a certain disease. In general language, the public understanding of *vaccine* does not differ significantly from that of medical experts, although the level of understanding may vary from person to person depending on their experience. Some may have a more in-depth understanding of vaccines, while others may have a fuzzier perception of the Term. In the majority of countries, our first encounter with vaccination happens in early childhood, at an age when we are not conscious enough to be aware of what a vaccine is. For instance, the vaccine against hepatitis B is recommended for newborns within first 24 hours of life. After that, we get vaccinated many times throughout our lives, making the process of being vaccinated, as well as the Term *vaccine*, a common part of our lives.

Let us consider two examples that illustrate the usage of *vaccine* in two different types of texts: example (8) is a piece of scientific discourse from the prestigious medical journal *Lancet*; example (9) is an excerpt from an ordinary speech given by a Minnesota resident.

(8) Other COVID-19 **vaccines** employ different adenovirus serotypes – the single-dose **vaccine** from Janssen uses an HAdV-26 vector and the **vaccine** from AstraZeneca uses the same chimpanzee-derived adenovirus vector for each of the two doses.
[Web, https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(22)00332-4/fulltext, 19/05/2022]

(9) The first time I got my **vaccine** I didn't take any pain reliever medication, but... when I took the second dose I took some Tylenol and that helped.
[Web, https://www.mnkaren.org/news/why-i-got-vaccinated/, 19/05/2022]

The examples given show *vaccine* is used in texts that belong to different registers. Example (8) uses technical medical vocabulary, such as *adenovirus*, *serotype*, *HAdV-26* (human adenovirus type 26) and *vector*, while example (9) is more informal and belongs to non-specialized discourse. However, in both cases, the meaning of the Term *vaccine* remains the same: a substance that is injected into the body to make it resistant to a disease. While the general public may have a more simplistic understanding of vaccines compared to medical experts, they still understand the basic procedure and objective of vaccination. Therefore, the terminological sense of *vaccine* is preserved in general language.

Another example of a runaway Term is the legal Term *mortgage*. Although *mortgage* has various senses in legal terminology, two primary senses are relevant here: (i) MORTGAGE **1** as a legal agreement for the purchase of a property, and (ii) MORTGAGE **2** as the amount of money borrowed by a borrower according to the terms of the mortgage agreement, that is, MORTGAGE **1**. In the modern world, millions of people have mortgage loans. As a result, the Term *mortgage* has become part of general language. Before signing a mortgage agreement, a borrower must understand the terms, which requires some legal and financial knowledge but not necessarily an extensive background in these areas. Therefore, the usage of *mortgage* in non-specialized discourse can still be somewhat informed.

Since a mortgage is often viewed as a financial burden, the Term *mortgage* is often used in negative contexts, as in both examples provided where *mortgage* corresponds to the terminological sense MORTGAGE **2**:

(10) My wife and I built our dream home on a nice plot of land and have been paying down the **mortgage** like mad.
[Web, https://www.financialsamurai.com/the-biggest-downside-to-paying-off-your-mortgage/, 29/05/2022]

(11) So I stopped having fancy holidays and nice cars and stuff and paid off the damn **mortgage** and built me a little stash of cash.
[Web, https://forums.theregister.com/post/reply/3058734, 28/06/2022]

In example (10), the author uses the intensifier *like mad*, which clearly implies how difficult it is to pay off the mortgage. In example (11), the author uses the pejorative adjective *damn*, which also contributes to the negative connotation of the Term. Despite the derogatory language used, we do not typically observe a distortion in the meaning of *mortgage* as both the legal Term and its layman counterpart refer to the same concept – the loan for the purchase of a real property.

In some cases, negative connotations may even lead to the emergence of a new metaphoric sense. In the Russian language, we are currently witnessing the emergence of a new sense of the Term *unomeka* 'mortgage loan'. This sense denotes a person's social and financial obligations that are considered by this person as a burden in life. What is more interesting, such obligations do not necesserily involve a mortgage obligation. We assume that this sense has resulted from the image of mortgage as a financial burden which partly deprives one of their liberty. Nevertheless, this sense is still too marginal and vague to provide a clear-cut definition.

Lexicographic description. Just like full Terms, runaway Terms have to be included in specialized dictionaries. However, in general dictionaries, runaway Terms have a better chance of being included in the nomenclature than full Terms as the latter are too specific for the general public. For example, the *Longman Dictionary of Contemporary English*, does not include *phytoplankton* in its nomenclature, whereas *vaccine* and *mortgage* are included.

To sum up, when used in speech production, the usage of a runaway Term may be twofold: (i) it may be used as a full-fledged domain-specific Term or (ii) it may be used as a simplified lay version in the general language context. Despite its twofold usage, we do not consider it necessary to split a runaway Term into two separate senses in a lexicographic source, that is, one for technical use and another one for laymen needs. This decision is based on close semantic similarity of both "senses." In the English Lexical Network, we propose to label such Terms with the usage note **(spec)**.

3.3.3 Quasi-Terms: **quasi-spec**

Characterization. As explained previously, some Terms become runaway Terms when they are picked up by general language while maintaining their original meaning, while others undergo significant semantic and grammatical changes and can no longer be considered Terms. In this case, a Term becomes a quasi-Term. Quasi-Terms are lay analogs of domain-specific Terms whose usage in general language discourse does not conform to the definition of their specialized counterparts. Quasi-Terms arise due to a lack of domain-specific knowledge and unconscious erroneous or loose interpretations of the terminological sense. From the Speaker's perspective, they may still believe they are referring to scientific concepts. Thus, new general language senses emerge.

First, we will discuss the notion of quasi-Terms in more detail. Second, we will contrast quasi-Terms with metaphoric uses of Terms.

Illustration. Take, for example, the linguistic Term *semantics*. In linguistics, we distinguish between three senses of SEMANTICS:

- SEMANTICS **1a** as semantic properties of a linguistic unit – *semantics of English prefixes*;
- SEMANTICS **1b** of a given language as a set of all the senses which can be expressed in this language – *semantics of Mandarin*;
- SEMANTICS **2** as a branch of linguistics that studies SEMANTICS **1a** and SEMANTICS **1b**.

In general language, the use of *semantics* often boils down to several senses that do not align with the three previously mentioned ones. Let us examine how the following examples illustrate the usage of *semantics* in non-specialized discourse:

(12) In recent years, traders have had to monkey around with the **semantics** when marketing their products. For example, they don't sell very small Mars bars, they sell you "fun-sized" Mars bars.
[Web, https://www.sunderlandecho.com/news/richard-ord-very-big-issue-shrinking-choccies-379644, 28/05/2022]

(13) And yes, I'm pretty sure that kind of nit-picking about specific wording and what they mean is indeed an argument over **semantics**.
[Web, https://countercurrents.org/2021/01/we-have-had-enough-of-answering-no-lessons-needed-on-composite-culture/, 28/05/2022]

(14) Let's stop playing with the **semantics** guys. Being prosecuted, alone, does not make you a convict. You have to be prosecuted AND convicted to become a convict.
[Web, http://www.city-data.com/forum/photography/1455673-illegal-get-pictures-developed-show-adult-3.html, 28/05/2022]

In (12), *semantics* can be interpreted as *wording*, i.e., the choice of words to express a certain idea. In this particular case, the choice of the name *fun-sized bars* instead of *very small bars* is called *monkeying aroung with the semantics*, i.e., using a euphemistic wording to avoid eliciting a negative response from consumers. In example (13), the author refers to *semantics* as both wording and meaning behind words. Finally, the author of example (14) uses the phrase *let's stop playing with the semantics* to suggest stopping speculating on the definition of *convict* and proposes their own understanding of the expression. As can be seen in these examples, the usage of *semantics* revolves around two main points: meaning and wording. Neither of the linguistic Terms – SEMANTICS **1a**, SEMANTICS **1b**, SEMANTICS **2** – can denote wording, making this usage a clear quasi-Term. However, *semantics* does refer to meaning. Thus, the author of (14) may not be entirely wrong in using *semantics* in the given context. However, the implicit message behind the sentence *let's stop playing with the semantics* is a request to stop discussing the language itself and to go back to the original discussion.

In fact, phrases such as *it is a question of semantics*, *it is a matter of semantics*, *argue over semantics*, *quibble over semantics* are frequently used in argument context, as in (13) and (14). Defining these phrases can be confusing as they allow for several interpretations. In general, when someone says *it's a question of semantics* in an argument, they usually imply (with a negative connotation) that the opponent's critique of another person's word choice is not significant and might not impact the outcome of the argument. Often, there is no significant conflict in the argument, and both parties agree on the same thing, but they are formulating their thoughts differently, leading to the use of such a phrase.

The general language material illustrating the use of *semantics* is vast, making it a tricky task for lexicographers to define all the developed general language senses. As seen from the above-mentioned examples, we can loosely define at least two senses: 'wording' as in (12) and 'unimportant details about the words as means of expressing content' as in (13) and (14).

Another example of a quasi-Term is the general language use of the word *antibiotic*. In a medical context, antibiotics are drugs that fight bacterial infections, such as skin, respiratory, sexually transmitted infections, and more. However, antibiotics are not effective against viral infections, like the common cold, flu, and chickenpox. The general public tends to perceive antibiotics as a universal remedy, even for minor illnesses like the common cold. This misconception stems from a lack of medical knowledge, as the general public usually does not differentiate between viral and bacterial infections. The uninformed use of *antibiotic* in general language transcends linguistic boundaries and affects how people actually take these drugs. Since antibiotics are perceived as a universal remedy, they are often used as one. Patients often expect the doctors to prescribe antibiotics regardless of the situation. In some countries, antibiotics are available for purchase without a medical prescription, which contributes to their uncontrolled use even against illnesses that antibiotics cannot fight. This can lead to even more serious health issues. Therefore, the medical Term *antibiotic* is distorted in general language, and the general language sense qualifies as a quasi-Term.

Additionally, consider the following examples of misused Terms from the general Russian language, as provided in Iomdin (2012), which align with our understanding of quasi-Terms:

- The anatomy Term *предплечье* 'forearm' refers to the part of the arm between the wrist and elbow, whereas in the general language it is commonly interpreted as the upper part of the arm between elbow and shoulder.
- The botany Term *мимоза* 'mimosa' that refers to a Brazilian plant with violet blossoms is often inaccurately used to refer to acacia tree with yellow blossoms.

Lexicographic description. When it comes to lexicographic description, it is up to lexicographers to decide whether a quasi-Term is persistent enough in general language to be included in the dictionary. However, defining quasi-Terms can be a challenging task, as their general language use often allows for multiple interpretations. If the decision is made to include the general language sense in the lexicographic source, it should be treated separately from the original terminological sense. In the English Lexical Network, quasi-Terms are labeled with the usage note **quasi-spec**.

Specialized dictionaries may not include quasi-Terms, while general dictionaries may label them with optional usage notes like **informal** or **colloquial**. It is important to note that quasi-Terms are not always recognized as full-fledged lexical units, and as a result, may be disregarded even in general dictionaries. We believe it is vital for dictionaries to not only describe established quasi-Terms, but also to contrast them with the full Terms they originated from.

Relation with metaphoric senses. We believe it is important to differentiate between metaphoric uses of Terms and quasi-Terms. General language senses based on metaphor and metonymy (*reboot* 'start something again') are established "legitimate" senses which do not claim to be specialized vocabulary. In contrast, distorted general language senses (*предлечье* 'the upper part of the arm between elbow and shoulder') are marginal cases of terminological misuse.

Take the chess Term *stalemate* as an example. In chess, *stalemate* refers to a situation where the player cannot make a move without being put in check. In general language, *stalemate* is used metaphorically to describe a situation where a problem seems unsolvable. The connection between the two senses is based on metaphor, mapping the desperate position on the chessboard to a real-life situation. When used in general language, the speaker is aware they are not using the original chess Term.

Similarly, consider the environmental Terms *climate* and *atmosphere* which are modified in general language into metaphoric *political climate*, *economic climate*, *friendly atmosphere*, *creative atmosphere*, etc. It's usually easy to differentiate between quasi-terms and metaphoric senses by analyzing their semantic link

to the original terminological sense.

However, in some cases, the distinction is more subtle. This occurs when a metaphoric general language sense is both based on metaphor and distorted intentionally by the Speaker. This is particularly evident in the case of medical Terms related to mental health conditions, such as *dementia*, *depression*, *OCD* and Terms used to describe individuals suffering from such diseases, such as *psychopath*, *schizophrenic*, *autistic*. While people are becoming more aware of the importance of mental health, there is a popularized usage of such medical Terms in general language. For instance, the term *dementia* is used in general language to refer to bizarre behaviour or illogical reasoning (15). *Autistic* is downgraded to refer to someone who is, according to the speaker's subjective judgement, stupid, antisocial or behaves weirdly (16). Similarly, the adjective *schizophrenic* is used to describe someone who changes their opinion too quickly and erratically (17).

(15) So when he says/tweets something incredibly dumb or acts stupidly, its gotta be **dementia**. Or Russia. Or he is distracting from something.
[Web, https://www.resetera.com/threads/trump-stop-human-hijacking.94117/page-3, 15/06/2022]

(16) Are you so **autistic** that you think that the camera on an iPhone and an actual camera are in any way comparable?
[Web, https://warosu.org/fa/thread/13751032, 2/07/2022]

(17) You're not actually making your end goal in any of your argument clear, and you're so **schizophrenic** in jumping from one argument to another that nobody can keep up.
[Web, https://gamefaqs.gamespot.com/boards/916373-pc/77986656 page=25#252, 15/06/2022]

It's clear that the speakers using these expressions are aware they are not using them in their original medical sense. Their message is metaphoric, and it can be interpreted as *I know you do not have dementia (autism, schizophrenia) but your behaviour makes me want to compare it to that of an ill person.* However, it is still of little relevance to the actual mental conditions as the speaker refers to the real symptoms of these conditions by exaggerating another person's behaviour or personal traits.

In summary, we argue that quasi-Terms is a specific type of general language senses that develops during domestication process. Quasi-Terms can be contrasted

with other types of general language senses, such as metaphoric senses, which result from the metaphoric mapping of domain-specific knowledge onto the general language context. Our semantic analysis shows that quasi-Terms appear when speakers lack domain-specific knowledge and unintentionally distort the original sense, while metaphoric senses arise from mostly intentional metaphoric mapping.

3.4 Environmental discourse and its terminology

3.4.1 Linguistic perspective on environmental discourse

From a linguistic perspective, environmental discourse refers to written and spoken productions about the environment. It is closely related to the notion of environmental terminology which is a system of Terms used to communicate information related to the environment. Environmental discourse is characterized by the following features:

Polyphony. Environmental discourse is polyphonic because it manifests itself in different channels, genres and voices: "scientific reports and research papers, news media articles, political manifestos and speeches, NGO programmes, White Papers, blogs, social media discussions and individual personal stories" (Flottum 2014b). Thus, environmental discourse exists in several forms, e.g., scientific discourse, political discourse, media discourse, business discourse, and general public discourse.

Dynamism. Environmental science is a dynamic discipline that is developing fast. As a consequence, environmental discourse and its terminology also change over time: new Terms emerge while old ones become obsolete or get revised. Neologisms appear due to the need to name new concepts and as a result of lexical creativity. For example, consider the lexical creativity around the Term *carbon*: *carbon finance*, *low-carbon diet*, *carbon sinner*, etc. (Nerlich, Koteyko and Brown 2010). On the one hand, such dynamism makes environmental discourse and terminology richer and more diverse. On the other hand, it leads to general terminological confusion. For example, new Terms may be unclear for some people due to the lack of proper definitions, e.g., *climate finance* (IPCC 2023a), *downcycling* (Helbig *et al.* 2022).

Terminological turbulence. Terminological turbulence in environmental discourse refers to cases when environmental terminology fails to effectively perform its intended functions. While there are many ways to intepret terminological turbulence, we will focus on the following issues: terminological variation, inaccurate translation, and lack of consistency in definitions. In environmental discourse, the single concept can be labeled with multiple Terms, and the difference between them is not always clear. For example, *global warming* is also referred to as *climate change*, *global climate change*, *climate warming*, *global heating* or *atmospheric warming* (Cabezas-García and León-Araúz 2023). Such terminological variation in environmental discourse can lead to confusion and hinder effective communication. Translation of environmental terminology can also pose challenges, especially due to cultural and language differences. For example, Allain (2013) questions the French translation of *substainable development* (**Fr.** *développement durable*) and concludes that it may not be the best translation, but it has already become a commonly used expression. Finally, the lack of precision and consistency in definining environmental Terms is a persistent issue in the field. Even the most prominent environmental Terms are still being disputed. Let alone neologisms and less popular environmental Terms which do not make their ways to news headlines. Despite efforts to address this gap, there is currently no single terminological source that provides a systematic coverage of all environmental Terms. The issue of terminological turbulence is not exclusive to environmental terminology only and is typical for other domains as well. For instance, Quirion (2006) studied linguistic and cultural terminological variability through the prism of marketing and localization.

Interdisciplinarity. Environmental discourse draws on multiple disciplines to describe and address environmental issues. Such interdisciplinary approach results in the use of many Terms borrowed from other fields, such as *carbon* (chemistry), *bioremediation* (biology), *bioaccumulation* (toxicology), etc. Furthermore, new interface Terms are constantly emerging which are built on the concepts relevant to the environment and other fields. For example, the Terms *carbon tax* and *environmental audit* appeared on the interface of the environmental science and economics, *environmental justice* – envi-

ronmental science and law, *green infrastructure* – environmental science and urban planning.

Ideological framing. In the face of environmental challenges, environmental discourse is often used as a tool to evoke specific public sentiment or response. To achieve this, speakers often use ideological framing which include specific choice of language, metaphors and style to convey a certain idea. For example, consider how António Guterres, the Secretary-General of the UN, characterized the risk posed by climate change: "We are on a highway to climate hell with our foot on the accelerator" (*The Guardian* 2022). Such statements are commonly labeled as *alarmist*. Environmental discourse is also closely intertwined with politics. As asserted in Flottum (2014a), "climate narratives are situated between science and politics." The political will to manipulate the environmental debate and construct politically appealing narratives often leads to bias and pushes the science to the periphery.

3.4.2 Popularization of environmental terminology

Many environmental Terms have already become buzzwords, and they are everywhere around us – from the mass media and political campaigns to advertisments, food labels, travel tickets and banners. We hear and see these buzzwords so frequently that they have become part of our daily lexicon. However, simply encountering these Terms regularly can provide a false feeling of expertise and comprehension. In reality, having a Term in our passive vocabulary does not necessarily imply our ability to accurately define it and use it. For instance, while we may know that bioenergy is considered good for the environment, we may struggle to provide a definition or explain how it is generated and used, and why it is considered better than conventional types of energy.

Previous research has shown that environmental Terms often create confusion among the general public. In survey conducted by Fletcher and Downing (2011), the respondents demonstrated poor knowledge of the following environmental Terms: *alternative energy*, *biodiversity*, *embedded carbon*, *green energy*, *low carbon*, *sustainable*. On the contrary, *organic*, *recyclable*, *zero waste* were relatively well understood. When asked the definition of *renewable energy*, one of the respon-

dents gave a very peculiar answer: "We know what it means but we don't know what it means" (Fletcher and Downing 2011: 3). According to another survey conducted by the company *Smart Energy GB*[9] in 2021, the following environmental Terms were identified as those that people were most likely to misunderstand: *greenwashing*, *biomass*, *net zero*, *regenerative*, *offsetting carbon*, *carbon neutral*, *carbon footprint*, *microplastics*, *biodegradable*, *sustainability*.

Poor understanding of terminology can have significant implications beyond occasional misuse or difficulty to define Terms. For example, the inability to comprehend the difference between the Terms *climate* and *weather* may lead to a general lack of trust in scientific climate predictions. Similarly, the choice of Terms in the field of water recycling can influence public perceptions of risks associated with consuming recycled water. In water recycling, reused water can be called by two synonymous Terms: *recycled* or *reclaimed*. The research conducted by McClaran *et al.* (2020) revealed that reused water labeled with *recycled* had less negative risk perceptions compared to *reclaimed*, which had more negative associations and less chances to gain public acceptance. The authors concluded that the positive connotation of *recycling* influenced the public perception of *recycled water*. Similarly, the use of Terms such as *local* and *organic* on food labels can also affect consumers' purchasing decisions (Campbell *et al.* 2015).

Educators are well aware of the problem and are taking steps to clarify the common points of terminological confusion. There are many voices advocating for the correct usage of environmental terminology, from NASA to fashion magazines that publish articles aimed at debunking confusing environmental Terms. Let us break down a couple of environmental Terms that have received particular attention in educational resources.

Climate change vs. *global warming*. The Terms *climate change* and *global warming* are often used interchangeably, but they have distinct meanings. The latest IPCC glossary (IPCC 2023a) defines *global warming* as a long-term increase in the average global surface temperature. *Climate change* refers to the change in the state of the Earth's climate that persists for a long period, over decades or longer.

[9] https://www.smartenergygb.org/resource-centre/press-centre/eco-jargon-could-be-damaging-climate-change-efforts?docspage=3

While *global warming* refers to the change in surface temperature only, *climate change* refers to a broader range of environmental changes, such as sea level rise, extreme weather events, and changes in precipitation patterns. In other words, global warming is only a part of a larger environmental issue called *climate change*. At times, the preference of one Term over another is rooted in politics. According to Schuldt, Konrath and Schwarz (2011), American conservatives tend to prefer the Term *global warming* while liberals prefer *climate change*. The representatives of the conservative movement have even removed the Term *climate change* from federal websites to play down public interest and restrict access to information on climate change (Davenport 2018). In order to avoid further confusion, large institutions such as NASA and IPCC have chosen to emphasize the Term *climate change* over *global warming*.

Climate vs. *weather*. The opposition of *weather* and *climate* also creates confusion in people's minds. For example, the National Centers for Environmental Information offers an image (Figure 3.1) to clarify that *weather* denotes short-term fluctuations in conditions such as air temperature, humidity, and wind speed, while *climate* denotes long-term changes that occur in a particular area.

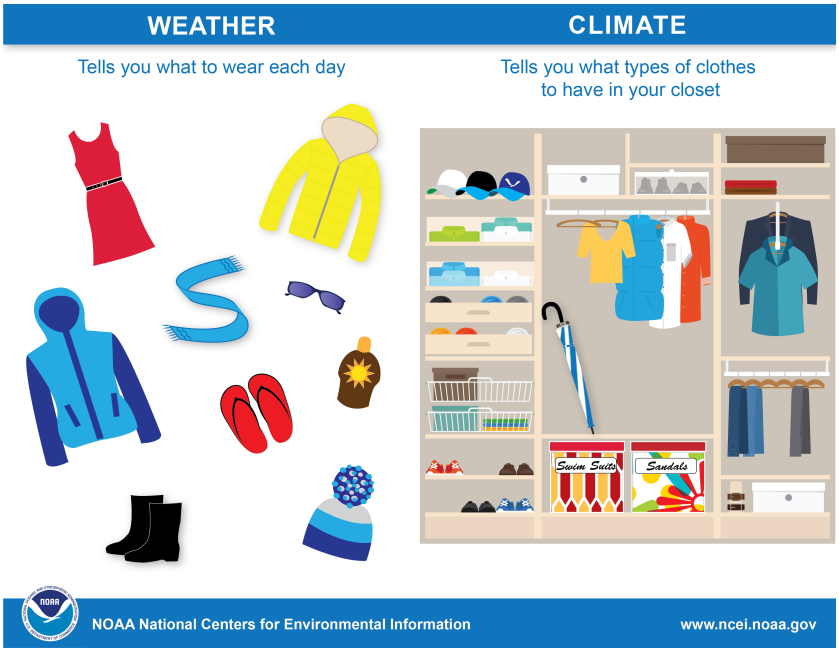

Figure 3.1 – NOAA: Explaining the difference between *weather* and *climate* with graphics.[10]

[10]Image source: https://www.ncei.noaa.gov/news/weather-vs-climate.

Organic. The polysemy of *organic* is quite extensive – the *Oxford English Dictionary* identifies 24 senses. In public discourse, *organic* describes natural foods grown without harmful chemicals like pesticides. However, this use of the Term is often ambiguous, misleading and serves marketing purposes, which has led educators to resent it (K. Chang 2008). The whole concept is often accused of being a matter of greenwashing. The Term *organic* is also used in chemistry to refer to any compound that contains carbon atoms bonded to hydrogen atoms, with some exceptions. The coexistence of these two Terms in environmental discourse may generate misunderstanding without proper instructions on their use. In one of the comment threads on the Web, the user complains: *I study the effects of organic acids on bacteria. Some people think that means that I do "health food" research* (K. Chang 2008). Even if we put aside the confusion between the two, *organic* as a label for natural foods is often misinterpreted, with many people mistakenly thinking that farmers do not use pesticides to grow organic food. In fact, they do. According to EPA (2022b), "pesticides derived from natural sources (such as biological pesticides) may be used in producing organically grown food." Ultimately, the label *organic* often gives consumers a green light for purchasing the product without the need for further investigation.

3.4.3 Terminology of current and emerging environmental issues

The topic of the environment is vast and complex, and analyzing linguistic data related to it can be overwhelming and produce inconclusive results. To perform a more focused and coherent analysis of environmental terminology, it is necessary to establish a specific perspective through which this terminology will be studied. Before arriving at our final decision to focus on the terminology of current and emerging environmental issues (which will be defined later in the present section), we tried several different perspectives.

Originally, we considered working on the "core" environmental terminology. In specialized language, the concept of *core* (also *fundamental*, *key*, *exemplar*) Terms often rely on corpus-dependent statistical measures, such as frequency, commonnes, and semantic criteria, such as universality. Mikhel (2022: 14) studied chemical terminology and established the nomenclature of *core*, or *fundamental*,

chemical Terms. These Terms represent the notional core of chemistry taught in general or introductory chemistry courses, e.g., *atom*, *bond*$_{(N)}$, *bond*$_{(V)}$, *element*, *molecule*, and *reactive*. Furthermore, this approach formalizes the hierarchy of chemistry Terms, with semantic fundamentals such as *substance* and *atom* (Mikhel 2022: 98) at its foundational level. Terms found in further levels (e.g., *molecule*, *element*) are semantically built based on semantic fundamentals and thus are more semantically complex. The formalization of such hierarchy is achievable in chemistry due to its relative homogeneity, and the Terms in this hierarchy represent the essential chemistry knowledge and concepts. In contrast, environmental science is a versatile discipline which borrows practices and terminologies from various fields, including chemistry. Hence, such hierarchization and identification of the most fundamental Terms in the entirety of environmental terminology appears non-feasible.

Furthermore, our research project stems from a long-term collaboration between a specialist in lexicology Alain Polguère (ATILF[11]) and a specialist in theoretical and green chemistry Francesca Ingrosso (LPCT[12]). This interdisciplinary research collaboration focuses on the linguistic aspects of chemical and environmental terminologies and has resulted in several publications on the subject, including Ingrosso and Polguère (2015), Mikhel (2022), Gotkova, Ingrosso, *et al.* (2023).

Working with a chemistry specialist provided the opportunity to explore the chemistry side of the environmental terminology. In 4.1, we will discuss the task of keyword extraction from specialized corpora related to both the environment and green chemistry. This extraction aimed to compile a list of environmental keywords pertinent to both environmental and chemistry discourse. However, as our thematic focus shifted to the terminology of current and emerging environmental issues, further work following the above-mentioned extraction task concentrated more on pure environmental texts. Nevertheless, the extracted keywords continued to serve as a reference. Moreover, our lexicographic work described in 5.5 included

[11] Analyse et Traitement Informatique de la Langue Française, **En.** Analysis and Computer Processing of the French Language

[12] Laboratoire de Physique et Chimie Théorique, **En.** Laboratory of Theretical Physics and Chemistry

chemistry-related vocables such as CARBON and ⌜CARBON DIOXIDE⌝.

Considering these points, we adapted our methodology multiple times. Our final decision in establishing thematic boundaries of the terminology under analysis was to focus on the **terminology of current and emerging environmental issues**. There is no single official list of environmental issues from a specific environmental institution, as different organizations may prioritize different issues based on their objectives and scope of work. However, many international organizations and institutions, such as the United Nations Environment Programme (UNEP), the Intergovernmental Panel on Climate Change (IPCC), and the World Health Organization (WHO), have identified and addressed various environmental issues through their reports, assessments, and policy recommendations. For example, some of the global environmental issues include climate change, loss of biodiversity, deforestation, air pollution, land degradation, waste management, etc. Additionally, many national and local environmental organizations have their own lists of environmental issues that are relevant to their specific agenda.

> Therefore, we can define *terminology of current and emerging environmental issues* as a set of the most significant and widespread environmental Terms used to communicate information about the risks associated with current and emerging global environmental issues, as well as information about environmental solutions and technical developments.

Furthermore, within this thematic scope, we established the four following lexicological criteria for a Term to be pertinent. These criteria partly align with the concept of *general environmental lexicon* outlined in Drouin, L'Homme and Robichaud (2018).

1. **Essential environmental concepts**: The Term refers to the most essential environmental concepts, such as *decarbonize*, *recycle*, *renewable energy*, *sustainable*, etc. Here, *essential* denotes indispensable and predominant concepts that shape the current environmental agenda. For example, the Term *renewable energy* refers to energy sources that are self-replenishing and have

a low carbon footprint, such as solar, wind, and hydroenergy. The use and promotion of renewable energy is a key solution in addressing global environmental issues.

2. **Pertinence to several subdomains:** The Term is pertinent to several environmental subdomains. For instance, the Term *pollution* is applicable to diverse subdomains like air pollution, water pollution, soil pollution, and radioactive pollution. However, we are equally interested in Terms that are more specific, such as *deforestation*, which is rather isolated within land use terminology instead of being an overarching concept.
3. **Semantic connections:** The Term exhibits strong semantic connections with other environment-related Terms. The Term *climate* is semantically connected with other environment-related Terms such as *anthropogenic*, *climate change*, *environment*, *weather*, etc.
4. **Presence in the general language discourse:** The Term exists on the interface of the specialized and general language discourse. It means that it is not exclusively specific to specialized environmental discourse and is diffused in more general language contexts like mass media texts, general public communication, education, scientific dissemination. This criterion emphasizes our interest in the interaction of the general public with the environmental terminology.

As Terms that satisfy these criteria are highly representative of environmental discourse, they can be used to search for meaningful discussions related to the environment on the Web. By searching for these Terms on social media platforms like *Twitter* and *Reddit*, we can find conversations related to environmental issues. For example, using search keywords like `species extinction`, `habitat loss`, `biodiversity loss` can retrieve public conversations about biodiversity since the corresponding Terms are widely used in discussions on this topic. In Chapter 4, we discuss the methodology of creating a list of such keywords and demonstrate how it can be used to search for environment-related discussions on social networks. This list of environmental keywords can be also useful for educational purposes, such as environmental science classes. Furthermore, it can also be used to create glossaries that facilitate environmental communication between

specialists and non-specialists, raising environmental awareness.

3.5 Chapter synthesis

In this chapter, we addressed two main points. First, we discussed the notions of *Term*, *terminology*, *determinologization*, *domestication*, *full Term*, *runaway Term*, and *quasi-Term*. Second, we examined the features of environmental discourse and its terminology, with a specific focus on the popularization of environmental terminology. While making scientific terminology accessible to a wider audience is beneficial for raising environmental awareness among the general public, it also entails terminological challenges, such as incorrect perception, misuse, and confusion among non-specialists.

Although engaging in terminological debates may appear futile in light of larger issues such as climate change, it is worth noting that the popularization of environmental terminology, its domestication, and misuse are all contributing factors to the confusion around environmental topic. Educating people about the nature of environmental issues is crucial, and emphasizing the importance of using proper Terms is equally essential. To analyze environmental terminology effectively, we established four lexicological criteria to narrow down our focus to the most significant and widespread environmental Terms, what we call the terminology of current and emerging environmental issues (3.4.3).

The following Chapter 4 will discuss the difference between the notions of Term and extraction keyword. Specifically, it will cover the building of a social network corpora using a list of environmental keywords.

In Chapter 5, we will present a detailed analysis of the notion of *carbon*. This analysis will be grounded in the earlier points we discussed, including the popularization of scientific terminology, domestication, misuse, public perception, and conceptualization.

Chapter 4

Building a social network corpus of environmental discussions

Summary

To access thematic data on social networks, particularly conversations related to environmental issues, we needed a comprehensive list of environmental keywords to perform the meaningful extraction. This chapter describes two interrelated tasks: the hybrid technique for compiling a list of environmental keywords using Natural Language Processing (NLP) tools and manual analysis, and the utilization of these keywords to retrieve online discussions from *Twitter* and *Reddit*.

In Section 4.1, we present an overview of a two-step process for extracting keywords from two specialized corpora: keyword extraction V1 and keyword extraction V2.

Subsequently, in Sections 4.2 and 4.3, we describe Twitter and Reddit, the extraction setup, and the filtering procedures applied to the obtained data.

Finally, in Section 4.4, we present the summary of two social networks corpora and offering a comparative analysis of the similarities and differences between the data obtained.

4.1 Identification of extraction keywords

4.1.1 Methodology

To build our thematic environmental corpora using data from social networks, we needed to compile a set of environmental keywords to extract relevant texts. The process of compiling the keyword list was not straightforward, starting as a manual compilation of a small seed list of 12 "core" environmental keywords (Table 4.1), which eventually expanded into the final version composed of 290 keywords (see Appendix A).

`atmosphere`	`greenhouse effect`
`carbon dioxide emission`	`greenhouse gas`
`carbon emission`	`ozone hole`
`climate change`	`ozone layer`
`environment`	`pollute`
`global warming`	`recycle`

Table 4.1 – Seed list of manually selected keywords.

To compile this list of 12 keywords, we posed ourselves a question: *Which environmental concepts come to mind first when we think about the environmental topic?* As a result, we built the list relying on our intuition and our existing knowledge of the environmental domain. The designation *seed* refers to the idea that this list served as a starting point towards a goal of compiling a more extensive list.

The quarantine-related constraints hampered the progress of automated tasks. Figure 4.1 provides a summary of the process which highlights the important milestones leading to the compilation of the final keyword list. We conducted two main iterations: keyword extraction Version 1 (V1) and keyword extraction Version 2 (V2) (discussed in 4.1.2 and 4.1.3 respectively). Throughout the entire keyword compilation process, we used hybrid methods, combining manual and automated selection.

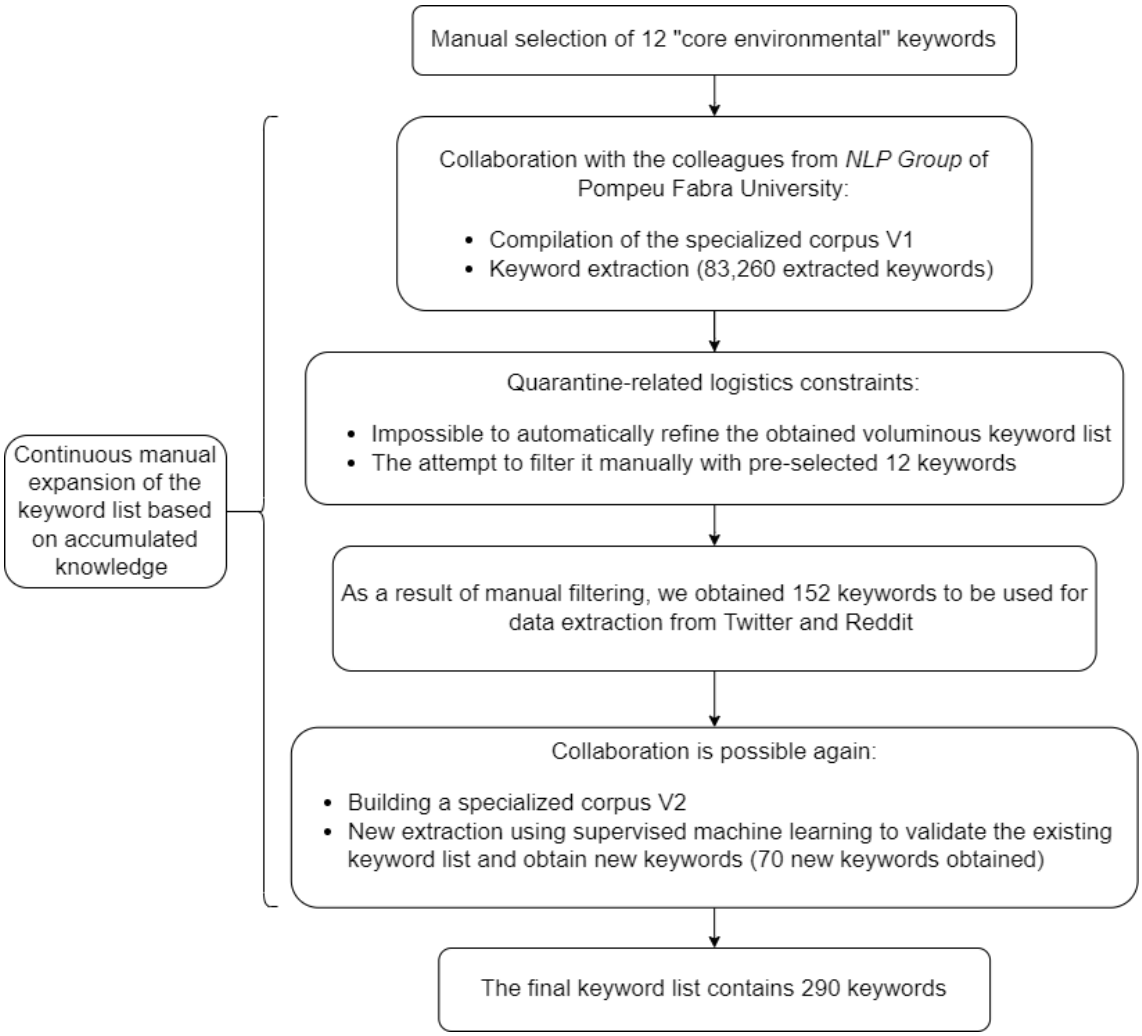

Figure 4.1 – The progress of compiling the keyword list.

Furtermore, as we accumulated knowledge, our conceptualization, terminology, and methodology for compiling the keyword list were changing. As discussed in 3.4.3, we reconsidered the notion of *core environmental Terms*, which later transformed into the notion of *terminology of current and emerging environmental issues*. Additionally, in keyword extraction V1, we placed equal emphasis on environment- and chemistry-related texts, whereas keyword extraction V2 prioritized the environmental discourse.

Notion of *extraction keyword*. Keyword extraction from a document or a corpus of documents is a well-known technique in NLP aimed at identifying the content that encapculates the text's "central themes, concepts, ideas, or arguments" (M. Q. Khan *et al.* 2022). In the literature, *keywords* are often conceptualized as meaningful words that express thematic centrality and statistical relevance to a given text. For example, Ahadh, Binish and Srinivasan (2021: 457) defined *keyword* as a content word that "contains information that contributes to the meaning of the sentence in which it occurs."

In the context of keyword extraction from specialized texts, keywords are equivalent to key Terms (Wu *et al.* 2020; Nomoto 2022), representing the most essential Terms that convey domain-specific knowledge. As stated by O. A. Khan *et al.* (2022), "keywords represent distinguished and specialized concepts and are bound to convey the informational content load of a document."

Our proposed approach to keyword extraction differentiates between keywords and lexical units, whether general or specialized.

> This differentiation highlights a fundamental distinction between them: a keyword is a semantically ambiguous linguistic form (signifier), whereas a lexical unit is always associated with a specific sense (signified). Therefore, Terms are objects of lexicographic description, while keywords are objects of automated extraction.

However, at the same time, keywords serve as proxies for Terms. This implies that even though a specific keyword is formally semantically empty, we still associate it with its respective environmental Term(s).

> We propose to define *extraction keyword* as a string of characters used for information extraction, with further tasks involving contextualizing these keywords within the discourse by identifying corresponding lexical units.

For example, the keyword `organic` (formatted in `teletype` font) on its own does not refer to any specific sense, but the respective lexical units acquire semantic features in context. When searching for the keyword `organic` (with no further specifications) in the COCA corpus, the following sentences were retrieved:

(1) Another team of scientists has produced a CO2-absorbing substance - one that binds the gas via a chemical reaction - by painting an **organic** compound called aziridine on a wafer of silica.
[COCA, Sid Perkins, 2008, Down With Carbon, Science News, Vol. 173, Iss. 16, pg. 18]

(2) Now, Lafranchi and his partner, Lynette Pareglio, grow **organic** tomatoes, strawberries, pumpkins and green vegetables.
[COCA, Jon Doyle, 2000, Land of Milk and Money, San Francisco Chronicle, pg. A13]

(3) Nevertheless, when diagnosing **organic** disease or psychiatric disorders, we should not understand distal or proximal causes as part of the malady itself.
[COCA, Valerie Gray Hardcastle, 1999, Multiplex vs. multiple selves: Distinguishing dissociative disorders, Monist, Vol. 82 Issue 4, p. 645]

In these sentences, *organic* corresponds to three different lexical units: ORGANIC **1** 'relating to carbon-containing chemical entity' (1), ORGANIC **2** 'relating to sustainable practices' (2), and ORGANIC **3** 'relating to body organ' (3).

4.1.2 Keyword extraction Version 1

Both stages of keyword extraction, V1 and V2, were done in collaboration with our colleagues from the *Natural Language Processing Research Group*[1] of the Pompeu Fabra University in Barcelona.[2] In the initial iteration, we used an extraction model described in Shvets and Wanner (2020) to extract a list of keyword candi-

[1] https://www.upf.edu/web/taln

[2] Leo Wanner helped us to outline the general strategy for keyword extraction process. Juan Soler Company contributed to gathering data for the specialized corpus V1. Alexander Shvets performed NLP-related tasks and fine-tuned the extraction model in both iterations.

dates[3] from two specialized subcorpora on the environment and green chemistry. The original *CE-PNG* model was trained on Wikipedia articles, specifically on highlighted "concepts" within those articles. Hence, the model is designed to identify complex concepts, denoted by nominal groups, which designate an entity in specialized domains, such as *Roman concrete*, *lime mortar*, and *triumphal arc*, as seen in the Wikipedia article on Roman architecture. In our application, we were equally interested in both simple and complex keywords. Although the extraction mechanism was adapted accordingly, it is worth noting that the model was primarily trained to extract nominal groups.

To apply the model to specialized textual data, we compiled two specialized subcorpora that satisfied the following criteria: 1) they consisted of research papers, official reports, news articles, and scientific blogs related to the topics of the environment and green chemistry, all written in English; 2) they were open access materials; 3) they were published no earlier than 2010; 4) they were issued by official organizations with expertise in the relevant fields. Given the scarcity of specialized resources on green chemistry, we also considered scientific blogs as a valuable source. The final version of the specialized corpus comprised a total of 17 million tokens, and the sources we used are listed in Table 4.2.

The extraction process had several key steps. First, the model was used to extract keyword candidates from two specialized subcorpora, which resulted in a total of 506,000 keyword candidates from the Environment subcorpus and 442,216 keyword candidates from the Green chemistry subcorpus. Furthermore, the two extracted lists were further filtered to obtain a merged list of 83,260 keyword candidates pertinent to both subcorpora. To filter out highly specific keywords, the merged list of keyword candidates was compared to the general lexicon units extracted from a general language corpus the *English Gigaword corpus* (Graff *et al.* 2003). This comparison allowed us to arrange the keyword candidates according to their specificity in our specialized corpus using the statistical measure *term frequency-inversed document frequency* (TF-IDF). Higher TF-IDF values indicated that a keyword candidate was more specific to our corpus compared to its distribu-

[3]Keyword candidates represent the material from which relevant keywords can be manually selected.

Source	Size	Period
Environment		
Scientific journal *Nature*	3,159 news articles 2.5M tokens	01/2010 - 02/2020
Scientific journal *Nature*	1,691 scientific papers 6M tokens	01/2010 - 02/2020
European Environmental Agency	866 news articles 400K tokens	01/2010 - 02/2020
Green chemistry		
Scientific journal *Nature*	1,831 scientific papers 7.4M tokens	01/2010 - 02/2020
Green Chemistry Nexus Blog by the American Chemistry Society	619 blog posts 300K tokens	12/2012 - 02/2020
Global Chemicals Outlook by the UN Environment program	2 reports 300K tokens	2013, 2019

Table 4.2 – Sources used for the specialized corpus V1.

tion in the general language corpus Gigaword. We performed a detailed analysis of keyword candidates and established a TF-IDF distribution range of scores from 50 to 70 to narrow down the search area to 3,833 keyword candidates. Furthermore, to obtain a final list of environmental keywords, we performed a manual analysis of 3,833 keyword candidates. As a result, we compiled a shortlist of 152 environmental keywords (see abridged version in Table 4.3). This list was further used to build our thematic environmental corpora using data from Twitter and Reddit. Finally, the result was validated by Alain Polguère (a specialist in lexicology) and Francesca Ingrosso (a specialist in green chemistry) to ensure its linguistic accuracy and relevance to the field.

4.1.3 Keyword extraction Version 2

We set two main goals for the second iteration of keyword extraction, referred to as "keyword extraction V2." Firstly, we aimed to validate the existing list of keywords, considering the shift in focus from the environment as a whole and green chemistry towards the topic of current and emerging environmental issues. It is important to note that during the transition from keyword extraction V1 to V2, we continued our study of environmental materials and complemented the keyword list obtained from keyword extraction V1 (see Table 4.3) by manually selecting

acidification	carbon credit
aerosol	carbon cycle
afforestation	carbon dioxide
air pollution	carbon emission
air temperature	carbon exchange
anthropogenic climate	carbon flux
anthropogenic warming	carbon footprint
aquatic ecosystem	carbon free
arctic warming	carbon neutral
atmosphere	carbon sequestration
atmospheric carbon dioxide	carbon sink
atmospheric methane	carbon source
atmospheric pollution	circular economy
atmospheric temperature	clean energy
atmospheric warming	climate
biodiversity	climate change
biodiversity hotspot	climate event
biodiversity loss	climate mitigation
biomass	climate model
bioshere	climate policy
black carbon	climate science
carbon	climate system
carbon balance	climate warming
carbon capture	co2
carbon concentration	coal

Table 4.3 – First 50 keywords of a bigger list of 152 keywords used to extract data for social network corpora.

additional keywords. This resulted in the updated keyword list of 268 keywords. Secondly, we wanted to obtain new keywords pertinent to the topic of current and emerging environmental issues.

The change in topic required a new specialized corpus, and we used this opportunity to be more selective in choosing sources based on their genre, topic

homogeneity, and recency. We focused on recent reports issued by large international authorities which actively address current and emerging environmental issues. These authorities included the *European Environment Agency*, the *Intergovernmental Panel on Climate Change*, the *United Nations Environmental Program* and the *World Meteorological Organization*. Below, we provide brief descriptions of each:

- The European Environment Agency (EEA) is an agency of the European Union founded in 1994 with the objective of providing knowledge and data about Europe's environment and climate. EEA provides assessments and informational support on various topics, including the state of Europe's environment, nature, health, climate, economy, resources, and sustainability.[4]
- The Intergovernmental Panel on Climate Change (IPCC) is a United Nations organization created in 1988. Its primary responsibility lies in evaluating the science of climate change by producing Assessment Reports that cover scientific, technical and socio-economic aspects of climate change.[5]
- The United Nations Environmental Program (UNEP), established in 1972, is an international environmental body within the United Nations organization. UNEP is responsible for the environmental agenda as it addresses current environmental issues by collaborating with governments, companies, and activists.[6]
- The World Meteorological Organization (WMO) is an intergovernmental agency of the United Nations created in 1950. It is in charge of international collaboration concerning the Earth's atmosphere, weather, and climate.[7]

The reports selected for the specialized corpus V2 address a range of environmental issues, such as climate change, biodiversity, circular economy. Table 4.4 provides an overview of the sources used to compile the corpus comprising 500 thousand tokens. The provided descriptions are direct quotations extracted either from the respective reports or the dedicated Web pages of the sources.

[4]https://www.eea.europa.eu/en/topics

[5]https://www.ipcc.ch/

[6]https://www.unep.org/

[7]https://public.wmo.int/en/our-mandate/what-we-do

Name and year	Description
The European Environmental Agency	
Trends and projections in Europe (2021)	This report explores the historical trends, most recent progress and projected future paths towards mitigating climate change through reduced greenhouse gas emissions, renewable energy gains and improved energy efficiency.
Knowledge for Action: Empowering the transition to a sustainable Europe (2021)	The report looks at [emerging environmental] issues from a range of perspectives: first, those relating to key policy areas — biodiversity and ecosystems; climate change mitigation and adaptation; circular economy and resource use; and human health and the environment. Secondly, a broader perspective is taken on sustainability trends, prospects and responses.
The Intergovernmental Panel on Climate Change	
Climate Change 2022: Mitigation of Climate Change (2022)	The Working Group III report provides an updated global assessment of climate change mitigation progress and pledges, and examines the sources of global emissions. It explains developments in emission reduction and mitigation efforts, assessing the impact of national climate pledges in relation to long-term emissions goals. [https://www.ipcc.ch/report/ar6/wg3/]
Global Warming of 1.5 °C (2018)	An IPCC special report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. [https://www.ipcc.ch/sr15/]
The United Nations Environmental Program	
The Sustainable Development Goals Report (2022)	With the COVID-19 pandemic in its third year, the war in Ukraine is exacerbating food, energy, humanitarian and refugee crises – all against the background of a full-fledged climate emergency. Using current data, The Sustainable Development Goals Report 2022 provides evidence of the destructive impacts of these crises on the achievement of the Sustainable Development Goals (SDGs).
UN Environment 2018 Annual Report (2018)	UN Environment released its 2018 Annual Report, highlighting the organization's work on issues from fighting pollution of the air and sea to helping nations meet their goals of reducing greenhouse gas emissions. [https://www.unep.org/resources/un-environment-2018-annual-report]
The World Meteorological Organization	
State of the Global Climate (2021)	The present WMO report provides an update on the annual state of the climate observed in the year 2021, and shows continued trends (also reported in the IPCC reports) in terms of key indicators. These include concentrations of greenhouse gases, global annual mean surface temperature, global mean sea level, ocean heat content, ocean acidification, sea-ice extent and changes in mass of the ice sheets and glaciers.

Table 4.4 – Examples of sources used to compile the specialized corpus V2.

Considering the limitations of the model used in keyword extraction V1, we conducted keyword extraction V2 through a comparative analysis of several models: *CE-PNG* and *T5*, namely *T5-small* and *T5-large*. T5, known as "Text-to-Text Transfer Transformer", serves various linguistic tasks, such as translation, summarization, keyword extraction, etc. (Raffel *et al.* 2020).

To create the trainig sample, we extracted sentences from the specialized corpus V2 containing 268 keywords which corresponded to 30,000 sentences out of around 50 thousand. Out of the 268 searched keywords, only 13 were not detected, such as `safe chemical`, `anthropogenic chemical`, `non-toxic chemical`, `benign chemical`, `green policy`, etc. As can be seen, these undetected keywords seem more relevant to the chemical aspect of the environmental topic and may be too niche for the subject of current and emerging environmental issues. Nevertheless, this addresses our first goal and confirms that the prevailing majority (95%) of preselected keywords are relevant to the topic of current and emerging environmental issues.

The training sample of 30,000 sentences was further divided into the training set, the development set, and the test set. The training set was used to teach the model, the development set was used to monitor its progress during training, and the test set was used to evaluate how well the model could find new keywords it had never seen before. We randomly shuffled the examples multiple times to ensure that each subset received a fair distribution of keywords in each. Specifically, 80% of the keywords were used for training, and these same keywords also appeared in the other two subsets. 10% of keywords were exclusively used in the development set, and another 10% were used solely in the test set, without overlapping with the training set. We kept this distinct 20% of keywords to test the model's ability to identify "unseen" keywords that it had not encountered during the training process.

To train the models, 268 keywords were divided into two sets: "core" and "supplementary" keywords. As mentioned earlier, in the course of our research, our methodology and terminology underwent several modifications, which also affected the division of the keyword list into "core" and "supplemetary" keywords as suggested in our original experiment described in Gotkova and Shvets (2023). Presently, we would perform such a division with meaningful lexical units and not with semantically empty keywords.

Eventually, tests showed that T5-large model outperformed other models in detecting "unseen" keywords, identifying 50-70% of them in a given set (62% across all the evaluation sets). Additionally, tests revealed that the inclusion of supplementary keywords into training enhanced the model's ability to detect core key-

words. Concerning the extraction of new keywords, we manually examined 171 non-annotated keywords extracted from the development set using T5-large. We concluded the analysis with the following statistics: 70 new keywords (41%), 32 keywords (19%) that already existed but were not automatically annotated due to parser mistakes, and 69 false negatives (40%) that were not keywords. Among the 70 novel keywords, 45 were combinations of existing keywords from our lists (e.g., `ecological drought`, `biomass contaminant`), while the remaining 25 keywords were completely new to our keyword list (e.g., `smog`, `renewable electricity`, `biomethane`). Although some these newly discovered keywords were quite specific (e.g., `cryosphere`), they were still considered important and valuable additions to our keyword list.

The final keyword list (see Appendix A) consists of 290 keywords resulting from keyword extractions V1 and V2, as well as continuous manual sampling. The list contains both simple keywords (`climate`, `biosphere`) and complex keywords (`carbon exchange`, `ecological degradation`). To avoid redundancy, we prioritized the base keyword[8] when dealing with keyword families. For example, within the keyword family of `acid`, `acidic`, `acidify`, `acidification`, we prioritized the keyword `acidification` for our list. This is because `acidification` is the most emblematic form and is more likely to appear frequently in environmental discourse, specifically in the context of ocean acidification, compared to its verb, adjective and noun counterparts. Clearly, when using the keyword list for further information extraction, it is advisable to use lemmatization to reach all the possible forms of a keyword used in texts, including plural forms of the listed nouns.

While we believe that our keyword list covers a significant portion of the discussions around environmental issues, it is important to acknowledge that this list is by no means exhaustive. The original intention behind compiling this keyword list was motivated by the need to use it for data extraction from social networks, a task that we performed using the intermediate keyword list resulting from keyword extraction V1. The final keyword list found in Appendix A was developed at a

[8]The notions of base keyword and keyword family are similar to the notions of base word and word family used in Bauer and Nation (1993): "a word family consists of a base word and all its derived and inflected forms," e.g., *watch*, *watches*, *watched*, *watching*.

later stage of our research.

It must be noted that when performing non-specialized and non-thematic data extraction using our keywords, particularly from social networks, it is important to distinguish between *autonomous* and *non-autonomous keywords*. Autonomous keywords are self-sufficient entities that can be used independently to extract environment-related information, e.g., `biodiversity`, `climate change`, `greenhouse gas`, etc. In contrast, non-autonomous keywords require additional complementary keywords to retrieve relevant texts. For example, the keyword `green` is non-autonomous because the corresponding adjective *green* is highly polysemous and not primarily linked to the environment. Thus, when extracting information, the keyword `green` should be accompanied by other keywords from the list. For instance, the extraction condition should indicate the combination of the keywords `green` and `pollution` in the same sentence. In Appendix A, such non-autonomous keywords are marked with an asterisk.

4.2 Extraction of Twitter data

At the moment of compilation of our corpora in 2020, we pinpointed the following reasons to chose exactly Twitter and Reddit for data extraction and corpus compilation:

- large worldwide audience;
- publicly available data, officially approved for extraction;
- inherent organizational difference between the platforms that would make them complementary to each other for a qualitative analysis;
- informal and colloquial nature of the posted content;
- versatile demographics of users.

4.2.1 What is Twitter 2020 Edition?

In 2022, the American businessman Elon Musk purchased Twitter which resulted in drastic changes to platform's internal policies following this acquisition. The history of Twitter has been divided in two periods, before and after the purchase deal. The changes included laying off 50% of Twitter stuff, reactivating certain

popular banned accounts (including that of Donald Trump), and introducing new subscription options, among other modifications (Kleinman 2023). On July 24, 2023, Twitter announced a major rebranding with the change of its name and iconic bird logo to an "X" logo, and other interface modifications like renaming tweets to "x's."[9] Despite all these perturbations, Twitter has not yet lost its position as one of the leading worldwide platforms. However, these changes did make some users consider switching to other similar platforms.

As mentioned in 1.1, to maintain consistent language throughout the text, we continue to use the former name *Twitter* to refer to the platform X.

Twitter[10] is an American social network platform that was founded in 2006. Its simple interface, accessibility via smartphones and the Web, and the fact that it was initially free to use (paid subscriptions have been recently introduced) made it appealing to a wide audience around the world. As of 2022, Twitter had 368 million active monthly users (Statista 2023b), making it one of the most popular and biggest social networks ranked among the top 15 social networks by size.

The original idea behind Twitter was to alow users to share short personal messages or "status updates" about their life in SMS-like manner. These short messages became known as *tweets*. Figure 4.2 provides a Web view of a tweet. This tweet consists of the text with *hashtags*[11] #`climate`, #`ClimateChange`, #`holidayseason`, #`holidays`, #`Spain` and two pictures. It also includes metadata like the posting time and the number of people who viewed this tweet (11.5K).

[9] https://www.bbc.com/news/business-66284304

[10] https://twitter.com/

[11] Hashtag is a string of words preceded by the "#" symbol that is used to index and categorize keywords and themes on Twitter. For example, users can use the hashtag #`climatechange` to indicate that their tweets are relevant to the topic of climate change. Similarly, they can search for this hashtag to discover other tweets that are related to the same topic.

Figure 4.2 – An overview of a tweet.

The communication on Twitter is primarily based on follower-followee architecture, meaning users can follow each other to subscribe to one another's content. This way, tweets from accounts that a user follows appear in their constantly updated personalized feed, i.e., their homepage. This following relationship need not be reciprocal, and no double-sided confirmation is required. Aside from tweets, other forms of communication on Twitter include replying to tweets, sharing them, liking them, quoting them, sending direct messages, etc.

As Twitter had been gaining its popularity, it attracted businesses, celebrities, organizations, news agencies, politicians, and others seeking new channels to disseminate information, communicate with customers, and reach new audiences. In Figure 4.3, we provide a tweet from the official account of the EU Directorate-General for the Environment.

A part of the content on Twitter is generated by social bots, which are accounts controlled automatically by pre-programmed softwares rather than humans. Although bots can contribute to misinformation (Xu and Sasahara 2022), they make up a significant part of the content on Twitter. For example, a study

[12]Tweet ID: 1679816085225234432

Figure 4.3 – A tweet from the account *EU Environment*.[12]

by (Chen *et al.* 2021) revealed that bots contributed to 15.4% of discussions in 491,279 analyzed tweets on climate change.

The diversity of users makes it challenging to profile an average modern Twitter user without considering additional criteria. When studying Twitter users, researchers often focus on specific aspects or subjects. For example, Toupin, Millerand and Larivière (2022) categorized Twitter users who share research papers on climate change into eight categories: academia, communication, political, professional, personal, organization, bots, and publishers. To our knowledge, there have been no attempts to identify distinct audiences tweeting about environmental issues from a general perspective.

Like any digital platform, Twitter is dynamic and always changing to adapt to users' needs and current digital trends. These changes can range from internal and privacy policies to alterations in the platform's appearance and tweet length.

For example, the length of tweets has undergone several modifications throughout the Twitter's history. Initially, tweets were limited to 140 characters, the limit was further extended to 280 characters in 2017, and as of 2023, users with a paid subscription can write tweets up to 4,000 characters long.

According to reports, the owner of Twitter, Elon Musk, accepted government demands for censorship on several occasions.[13] Given this trend, and considering that the environmental topic is highly politicized, especially in the United States, it is likely that conversations about the environment on Twitter are more moderated as they were before the acquisition. Furthermore, a recent study conducted by C. H. Chang *et al.* (2023) revealed that half of the investigated 380,000 users with environmentally oriented profiles stopped their activity on Twitter since the sale deal in October 2022.

> Given these points, it is important to highlight that our Twitter corpus was compiled in its 2020 edition, in the "pre-Musk" era of Twitter.

Twitter data for academic purposes. Twitter data is widely used in research and has proven to be a valuable source for various academic projects. The richness of Twitter data demographics, in terms of age, gender, and geography, allows for both quantitative and qualitative studies. The application of Twitter data in research is truly diverse across numerous disciplines, and we will provide a couple examples pertinent to environmental topics.

Zarrabeitia-Bilbao *et al.* (2022) analyzed 400,000 tweets to explore the impact of the *United Nations World Environment Day* programs on public discourse. Kolic *et al.* (2022) investigated the retweet dynamics and patterns of Twitter discussions to identifiy echo chambers of climate believers and skeptics. Shi *et al.* (2020) examined diverse Twitter discourses related to environmental issues, with a specific focus on two competing hashtags #`climatechange` and #`globalwarming`. Additionally, S. C. Kim and Cooke (2018) studied the framing of emerging environmental issues, such as ocean acidification and climate change,

[13] https://edition.cnn.com/2023/05/29/tech/elon-musk-twitter-government-takedown/index.html

before and after the US President's announcement of the U.S. withdrawal from the International Paris Climate Agreement.

The ethics of using the "*Big Data*"[14] as an object of extraction, analysis and further dissemination in academic research has been extensively discussed in the scientific community. Twitter has provided solutions for data extraction, making it accessible to parties interested in obtaining large amounts of Twitter data through different types of Application Programming Interfaces (APIs), some of which allow access to the whole archive of tweets.

Due to the significant interest in Twitter data within the academic community, Twitter has introduced the *Twitter API for Academic Research* to facilitate the data extraction for researchers. Although our corpus was compiled before this innovation, we were able to extract the data using *Streaming API* by following the standard procedure which included creating a developer account and providing a detailed project description (see 4.2.2). In 2023, Twitter's API for academic research became a paid service with restricted access to data.

As of July 21, 2023, according to the current version of terms of use for developers (Twitter Developer Platform 2023), academic researchers are allowed to redistribute content by sharing Tweet IDs and User IDs only:

> If you provide Twitter Content to third parties, including downloadable datasets or via an API, you may only distribute Tweet IDs, Direct Message IDs, and/or User IDs (except as described below). [...] Academic researchers are permitted to distribute an unlimited number of Tweet IDs and/or User IDs if they are doing so on behalf of an academic institution and for the sole purpose of non-commercial research.

4.2.2 Data extraction setup

Data extraction from Twitter is quite straightforward due to Twitter's API and its clear official documentation on how to use it for data extraction. To gain access to the API, developers are required to register an application through the

[14] *Big Data* is a Term typically used to denote extremely voluminous and dynamic Web-based data like social media texts. Big Data has become a valuable source in computational and corpus linguistics.

Twitter Developer Platform. This involves filling out a detailed form with the project description and explaining the goals of using the extracted data. Once the application is approved and the developer account is created, a developer obtains personalized credentials to be able to connect to the API.

For our data extraction, we used the Twitter's *Streaming API*[15] that provides access to 1% of the "live" data as a random set of a global continuous stream of publicly available tweets. To connect to Streaming API, we wrote a Python code using the *Tweepy*[16] library. This code was set up to gather tweets related to the environmental topic. To ensure we obtained only relevant tweets from this random set, we established several filtering criteria:

1. The tweet must contain at least one keyword from our list (see 4.1.2).
2. It must be written in English.
3. It should not be a *retweet* to avoid populating data with duplicates.

Although the Streaming API offers the above-mentioned filtering options, they were not flexible enough for our needs. For example, it was not possible to use regular expressions[17] to automatically search for various forms of a keyword such as `pollute`, `polluting`, `polluted`. Therefore, we manually added these wordforms to our keyword list. Furthermore, it was not possible to use Boolean operators (AND, OR, etc.) to look for combinations of keywords to eliminate irrelevant texts. Searching for non-autonomous keyword `atmosphere` in combination with any other keyword from the list would help eliminate irrelevant contexts, such as *We enjoyed the friendly atmosphere.* However, we managed to address this during the text normalization stage.

The setup code ran continuously for 20 days (02.09.2020 - 21.09.2020) and stored the data in a *MongoDB* database in JSON format. We chose MongoDB as a NoSQL database for this task due to its efficiency and flexibility when working

[15]https://developer.twitter.com/en/docs/tutorials/consuming-streaming-data

[16]https://www.tweepy.org/

[17]"A regular expression is a pattern that the regular expression engine attempts to match in input text." [https://learn.microsoft.com/en-us/dotnet/standard/base-types/regular-expression-language-quick-reference]

with large volumes of continuous data from Streaming API. Each extracted tweet was accompanied with rich metadata comprising 112 information fields (see Figure 4.4). Ultimately, our raw Twitter dataset extracted in the timeframe of 20 days comprised a total of 4,709,014 tweets.

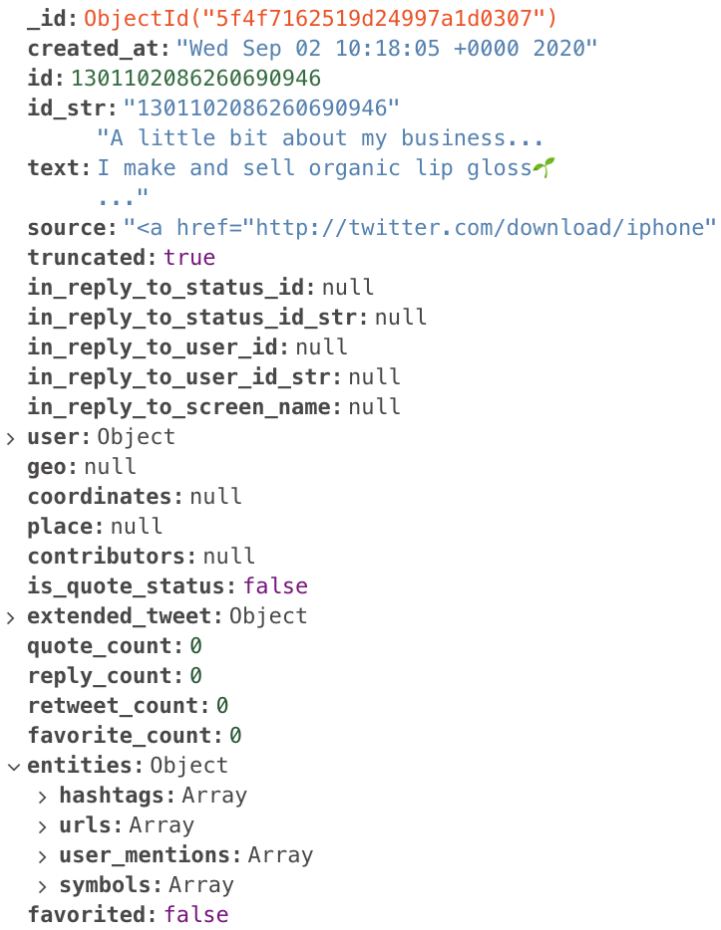

Figure 4.4 – An overview of tweet metadata.

4.2.3 Data normalization

Despite the filtering conditions we implemented in our Python code for data extraction, the extracted data still contained tweets without keywords and duplicate tweets. However, this is a common phenomenon since Twitter warns about the imprecision of the API filters. Consequently, we had to further process the database to remove such tweets. First, we deduplicated the data by deleting 79,637 duplicate tweets, which accounted for 1,7% of the entire dataset. Furthermore, to exclude irrelevant discussions, as in (4) and (5), and tweets with no keywords, we filtered the data against our keyword list again, while making sure that non-autonomous

keywords (see 4.1.3) are combined with other keywords. While this step resulted in the removal of a significant portion of tweets – 2,628,968 (57%), – we obtained a corpus of more meaningful environment-related discussions. Figure 4.5 provides a screenshot of the Twitter database stored in a .csv file.

(4) But even if I was a girl, I wouldn't fall for him. But the **atmosphere** during that moment was good. It was an embarassing scene but with a friendly **atmosphere**, the shooting ended well. [Twitter, ID: 1304044972019130372]

(5) The political **climate** now is one of anarchy and dictatorship. You must agree that dissenting voices and opposition could thrive Pre-2015 because it was a different **climate**. One, where the rule of law and Democracy thrives despite the many shortcomings of that regime. [Twitter, ID: 1301559381725773829]

created_at	user.id	id	text
Wed Sep 16 17:13:22	29426700	1306280027609329664	#WorldOzoneDay reminds us of the need to protect the ozone layer, which absorbs most of the sun's harmful ultraviolet light. The ozone layer helps protect humans from skin cancer and shields plants, crops and ecosystems from damage. https://t.co/MK4FyKvjtS via [redacted]
Tue Sep 15 23:06:31	1082124456271716352	1306006509852258304	How is arson related to climate change? Maybe the burning of our city caused the hazy skies! What did that have to do with climate change!?
Fri Sep 18 23:43:45 +	2345388363	1307103047143878657	I saw an article on Livescience that showed a newspaper article from I believe it was 1909 if I'm not mistaken... saying how (i'm paraphrasing) mass burning of fossil fuels generated by the Industrial Revolution could lead to an increased greenhouse effect and higher temps.
Wed Sep 09 12:16:53	2887280459	1303668700097249282	[redacted] It's there as I have said and in Green. The EU one I didn't mention as it says what I said that most US standards are voluntary. The US has lower standards & measures to meet environmental standards. There are chemicals ad nauseum. Across the board it's disgusting. Section 3
Sat Sep 12 22:43:32 ·	119838257	1304913564654866432	[redacted] Which Biden will show up? The one who will end fracking & fossil fuels or the one who is for fracking & fossil fuels?
Sun Sep 13 23:33:26	2886020041	1305288508144148480	i guess u should say climate change not global warming but u get the gist.
Sat Sep 19 11:26:19 ·	916769311	1307279853360562177	[redacted] Waste management in the times of # covid-19 has created a large environment al crisis due to lack of infra.the practices of burying, openly burning of waste are prevelent. This needs to be addressed on priority.
Thu Sep 10 08:58:10	726563436402860032	1303981079859089409	[redacted] This article isn't very good. Downplays the role of climate change, & makes recent past disasters (fires, hurricanes, etc) seem like an apocalypse. It would be better to give practical advice on how to manage it, the new normal. In other words, a boring waste of pixels & print.
Mon Sep 14 00:28:00	851599873	1305302241335808001	Climate change destroyed the Indus Valley civilization 3000 years ago. Of course their cars were not very pollution free

Figure 4.5 – A screenshot of Twitter database stored in a .csv file.

Furthermore, Figure 4.6 shows the top 20 most frequent keywords from our keyword list that occur in the filtered Twitter data.

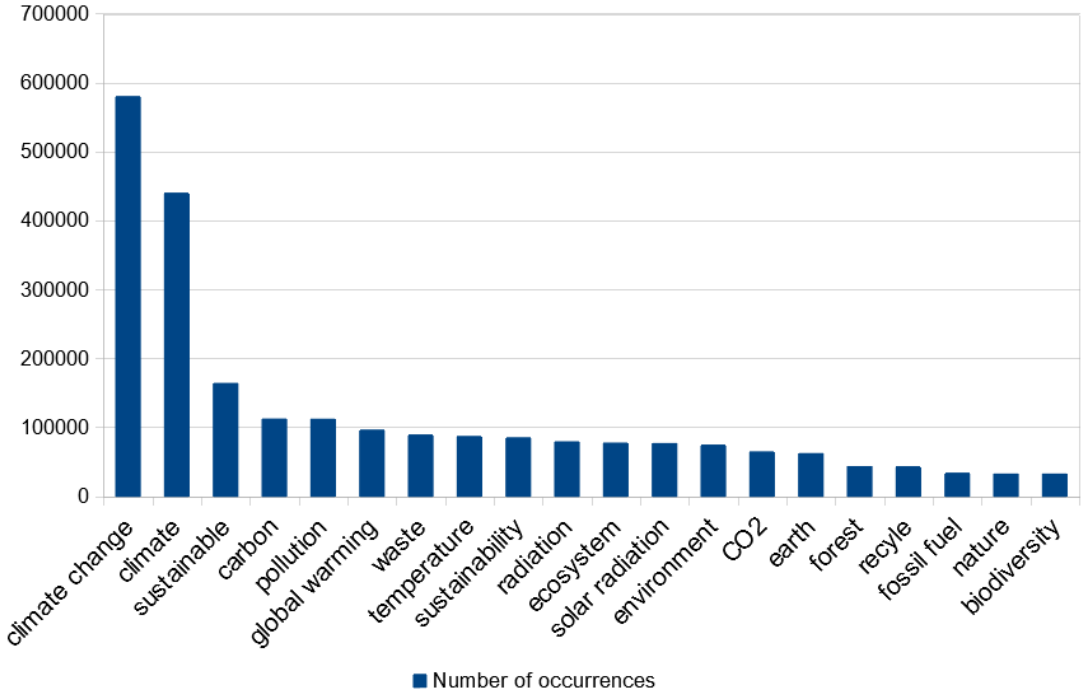

Figure 4.6 – Top 20 keywords that occur in the filtered Twitter data.

When preparing corpus data for NLP tasks, such as sentiment analysis, researchers prefer extensive data cleaning to ensure the removal of all the "noise." This includes the removal of URLs, hashtags, mentions, punctuation, stop words,[18] etc. As a result, the cleaned text may not resemble the original version, as it becomes a string of lemmatized tokens fed into the automated model to identify the positive or negative sentiment.

We performed this kind of cleaning to obtain statistical data about our corpus, such as the number of tokens. However, we chose to keep certain artifacts, such as user mentions, hashtags, and links for the qualitative analysis, as they contribute to the meaning of the text. For example in (6), the hashtag #`Methane` is an integral part of the sentence, and its removal would distort the meaning of the sentence. In contrast, the hashtags #`Climatecrisis` and #`ClimateEmergency` are added after the end of the sentence and serve as entitites that connect the tweet to the general pool of tweets on environmental issues, specifically the climate crisis and climate emergency.

(6) #**Methane** is a potent greenhouse gas, roughly 25 times more potent than carbon dioxide, and it is estimated that methane frozen in these sediments constitute the

[18] In NLP, stop words are mostly pronouns, articles and prepositions, e.g., *I*, *a*, *the*, *for*.

largest organic carbon reservoir on Earth. **#ClimateCrisis #ClimateEmergency** [Twitter, ID: 1301167005068820480]

4.3 Extraction of Reddit data

4.3.1 What is Reddit?

Reddit[19] is an American social network that was founded in 2005. As of 2022, Reddit had a monthly active audience of 430 million users (Statista 2023a). While Twitter communication is typically centered around "spontaneous opinions" (Guerris *et al.* 2020), Reddit is a forum-like platform.

Reddit promotes communication within thematic communities, known as *subreddits*. These subreddits are places where users with similar interests gather, with these interests covering a wide range from hobbies to specialized topics. For example, the subreddit r/`DIY` (short for *do it yourself*) has 22 million members who share interest in crafting. The subreddit r/`Physics` with 2 million members gather experts and enthusiasts in physics.

In Reddit, each subreddit is a collection of posts, known as *submissions*, that are generated by Reddit users. These submissions often trigger further discussions in the form of comments – texts where users express their opinions or make a statement related to the original submission. Additionally, users can vote up or down on each submission to show their interest or lack thereof in the topic. The body of a submission can include a text post, a link or be left empty. For example, submissions in the subreddit r/`environment` mainly consist of links to news articles (Figure 4.7). In contrast to Twitter's character limits, Reddit allows for more extensive posts and comments, with a limit of 40,000 characters for a post and 10,000 characters for a comment.

Each subreddit has its own set of rules to moderate the communication and promote better practices within a given subreddit. Such moderation contributes to the homegenity of the language, topics, and interests discussed. Furthermore, this ensures that the submission will stay within the designated theme, which is

[19] https://www.reddit.com/

Figure 4.7 – A submission from the subreddit r/`environment`.

beneficial for compiling a thematic corpus, as in our case. For example, here are some rules of the r/`environment`[20] subreddit:

1. Submissions must be in English.
2. No satire. No memes.
3. No old news, duplicate submissions or multiple submissions on the same subject.
4. No trolling, no misinformation – blatant disregard for accepted science is not allowed.

Given that Reddit hosts thematic communities related to specialized domains, it is reasonable to assume that these communities attract specialists in their respective domains. Reddit users are conscious about their privacy, so to our knowledge, there is no direct way to examine the demographics of a given subreddit, particularly in terms of the number of experts involved. However, it is common for Reddit users to share their professional expertise in relevant discussions, as seen in the following examples where authors explicitly reveal their identity: *as an environmental scientist* in (7), *I'm not a botanist* in (8), and *microbiologist here* in (9). However, it is important to remember that the content posted on social networks is not always verified, and any information posted can potentially be fictitious.

(7) I am not disagreeing with what you have stated but as an environmental scientist I can say you are focusing too much on one aspect of the benefits that trees bring. [Reddit, ID: f5wsr0i]

[20] https://www.reddit.com/r/environment/

(8) Now, I'm not a botanist, but since trees store carbon, doesn't it make sense to chop down fully grown trees and immediately replace them with saplings on a regular basis? It seems like fully grown trees would suck less carbon from the air than a faster growing sapling. [Reddit, ID: f1iemex]

(9) Microbiologist here. This is very interesting from a bacterial physiology perspective, but I don't quite understand the application. [Reddit, ID: f9dg97i]

The environmental topic is well-represented across numerous communities on Reddit. As of 2023, there are 156 communities related to the general environmental subreddit r/`environment`.[21] These communities range from large membership communities focused on aggregated environmental information, such as r/`climate` and r/`ecology`, to more specific ones, e.g., r/`composting`, r/`OrganicFarming`. In 4.3.2, dedicated to data extaction on Reddit, we provide a list of environmental subreddits chosen for the compilation of corpus along with their corresponding descriptions.

Reddit for academic purposes. Similar to Twitter data, the applications of Reddit data in research are diverse and extensive. Proferes *et al.* (2021) conducted a review of 727 academic documents that used Reddit data and found its major applications in disciplines such as Computer Science, Engineering and Math, Medicine and Health, Social Science, Humanities, and Natural Sciences. Additionally, Reddit data has been used for research on environmental topic, although not explicitly mentioned as a separate category in the study. For example, Treen *et al.* (2022) studied the polarization of climate change discussions on Reddit data from 2017. Shah, Seraj and Pennebaker (2021) explored Reddit discourse concerning to climate-related events, such as natural disasters, political and policy events, to understand their influence on online discussions. Furthermore, Sadeghian *et al.* (2022) focused on sustainable mobility and analyzed Reddit comments from environment-related and car-related subreddits to understand human mobility behaviors. Mkono and Hughes (2022) examined Reddit posts dedicated to sustainable travel and tourism, with a specific emphasis on the concept of "sustainability paralysis."

[21] https://www.reddit.com/r/environment/wiki/related_reddits/

Due to the growing demand for Reddit data, in 2023, Reddit is introducing strict restriction on data access for certain third-party applications and charging fees, especially for companies using Reddit data to train artificial intelligence models. While the updated Reddit's Terms of API[22] do not explicitly mention the terms of data redistribution, we contacted the Reddit team and received the following answer:

> Use for research purposes is OK provided you use it exclusively for academic (i.e. non-commercial) purposes, don't redistribute our data or any derivative products based on our data (e.g. models trained using Reddit data), credit Reddit and anonymize information in published results, and otherwise access and use our Data API in accordance with our Data API Terms and Developer Terms.

In this response, it is explicitly stated that redistributing Reddit data is not allowed. In view of this information, we assume it is safe to adopt the practice established by Twitter and share only the IDs of the posts we have collected. In this way, we avoid sharing any personalized data, and the interested parties can access the texts associated with these IDs through Reddit's API.

4.3.2 Data extraction setup

Reddit data was collected using the Reddit's official API wrapper *PRAW*.[23] The procedure of obtaining developer access to the data is similar to that of Twitter. To extract Reddit data using the API, it is necessary to register an application and obtain personalized credentials.

While Twitter data extraction was keyword-based, the Reddit data extraction consisted in a two-step process. First, we extracted content from thematic environmental communities, and then we filtered this data against our list of keywords. This distinction in extraction techniques was based on the differences in the structure of these two social networks and the functionality of their respective APIs. Twitter's architecture relies on feed of primarily unrelated tweets, whereas Reddit

[22] https://www.redditinc.com/policies/data-api-terms

[23] https://praw.readthedocs.io/en/stable/

is built around thematic subreddits, making it a community-based platform. We established the following criteria in selecting Reddit communities for data extraction by focusing on English-language communities related to the environmental topic with no fewer than 5,000 members. The fourteen subreddits chosen for data extraction are presented in Table 4.5.

Subreddit	**Created**	**Members**	**Description from "About community"**
r/environment	25.01.2008	1.5M	Current news, information and issues related to the environment.
r/ZeroWaste	20.02.2013	1.1M	We are responsible citizens who try to minimize our overall environmental impact.
r/climate	07.05.2008	164K	NA
r/RenewableEnergy	23.07.2010	118K	Share your fascinating links about renewable energy. Discuss new renewable technologies. [...]
r/ClimateActionPlan	19.08.2018	85.1K	An action-oriented subreddit highlighting the active measures being taken to combat/mitigate and or adapt to climate change. [...]
r/ecology	13.11.2008	76.9K	NA
r/ClimateOffensive	29.10.2018	67.4K	We're here to do something about climate change. We're not here to talk about why it's happening, how bad it is, or who to blame. [...]
r/environmental_ science	30.12.2010	48.4K	This subreddit is for the *scientific discussion* of topics in the environmental sciences, geosciences, and other relevant discipline's; including papers, articles, research, public-policy, and both educational and professional advice.
r/Green	25.01.2008	39.3K	This subreddit is for issues relating to Green issues, including (but not limited to) Green Politics.
r/StormComing	23.01.2013	23.1K	StormComing is about Exceptional World-Wide Weather and Extreme Events. [...]
r/recycling	12.10.2010	18.3K	A home for the 'green' Redditor. Discuss tips on recycling common household materials, post interesting links.
r/Sustainable	10.01.2009	18.1K	Sustainable energy, food, water, air, living. Things/actions that help the planet and us live together without turning it into a perpetual disaster or dead planet. [...]
r/enviroaction	23.02.2013	12.1K	A subreddit dedicated to Environmental *Action* - petitions, fundraisers, events and other actions related to education about and protection of the environment.
r/GlobalWarming	NA	NA	NA

Table 4.5 – The overview of subreddits chosen for data extraction.

To ensure the relevance of the data obtained, we focused on extracting "top of all time" submissions from each subreddit. These submissions, representing the most upvoted content, are considered highly representative of a given subreddit (Roussel, 2019). Furthermore, such submissions display high community engagement, expressed in the higher number of comments, which is essential for obtaining rich and meaningful data. Due to the limitations of the Reddit API, it was not possible to choose a precise timeframe for extraction. As a result, our Reddit corpus contains comments and submissions which date back to 2004. However, such older texts form a small minority of our overall corpus. It is also important to note that the Reddit API allows a maximum pull limit of 1,000 submissions per subreddit, but this constraint does not apply for comments.

Each submission was extracted with the following metadata: name of the subreddit it belongs to, submission title, submission ID, number of comments, date and time of creation, and submission text. As for the comments, the following metadata were extracted: comment ID, name of submission the comment belongs to, comment text, date and time of creation. Our raw dataset extracted from Reddit comprised 13,939 submissions and 430,540 comments within a time span from 17/09/2004 to 14/05/2021. Figure 4.8 provides an overview of a database stored in a .csv file.

topic	body	comm_id	url	author	timestamp
dp599n	This should be the taxation requirement on the billiona	f5sxl9n	/r/ClimateActionPlan/comments/dp5	[redacted]	2019-10-30 23:27:54
dp599n	To be honest, this project is more about sending a mes	f5v1t1p	/r/ClimateActionPlan/comments/dp5	[redacted]	2019-10-31 09:41:05
dp599n	Imagine how if Jeff Bezos or Bill Gates breathed on th	f5t87xg	/r/ClimateActionPlan/comments/dp5	[redacted]	2019-10-31 01:23:27
dp599n	Considering that those people's whealt comes from the	f5snbxf	/r/ClimateActionPlan/comments/dp5	[redacted]	2019-10-30 21:12:28
dp599n	Plant hemp over trees. Hemp starts to abrosb CO2 im	f5t04qt	/r/ClimateActionPlan/comments/dp5	[redacted]	2019-10-30 23:56:38
dp599n	Not taking away from Elon's contributions or overall ac	f5t1vhl	/r/ClimateActionPlan/comments/dp5	[redacted]	2019-10-31 00:15:58
dp599n	Thats like the average person donating two cents of $1	f5t19it	/r/ClimateActionPlan/comments/dp5	[redacted]	2019-10-31 00:09:10
dp599n	/r/LoTR Turns out that Elon Musk is a fan of the Ents!	f5t6zg1	/r/ClimateActionPlan/comments/dp5	[redacted]	2019-10-31 01:10:27
dp599n	Somebody already donated more than Elon lol	f5uptsh	/r/ClimateActionPlan/comments/dp5	[redacted]	2019-10-31 08:15:39

Figure 4.8 – A screenshot of Reddit database stored in a .csv file.

4.3.3 Data normalization

In the extracted Reddit data, we have two types of texts that require different processing: submissions and comments. As mentioned previously, a submission may consist only of a title or may be accompanied by additional text. We found that out of the 13,939 extracted submissions, only 1,450 (10%) were accompanied by text. The remaining 12,491 submissions without any text were excluded from further analysis because they lacked textual value.

Although our data was extracted from thematic communities on Reddit, where it can be assumed that all texts address to some extent the environmental topic, we filtered the data against our keyword list to align Twitter and Reddit corpora. As a result of this filtering process, 546 (38%) of the submissions and 339,401 (79%) of the comments were removed. Eventually, the cleaned data contained 903 submissions and 91,140 comments. Figure 4.9 displays the top twenty most frequent keywords found in the cleaned Reddit data from our keyword list. Remarkably, 14 keywords (`climate`, `climate change`, `waste`, `sustainable carbon`, `pollution`, `global warming`, `waste`, `environment`, `co2`, `earth`, `forest`, `recycle`, `fossil fuel`, `nature`) appear in both the top 20 keywords of Reddit data and Twitter data (cf. Figure 4.6).

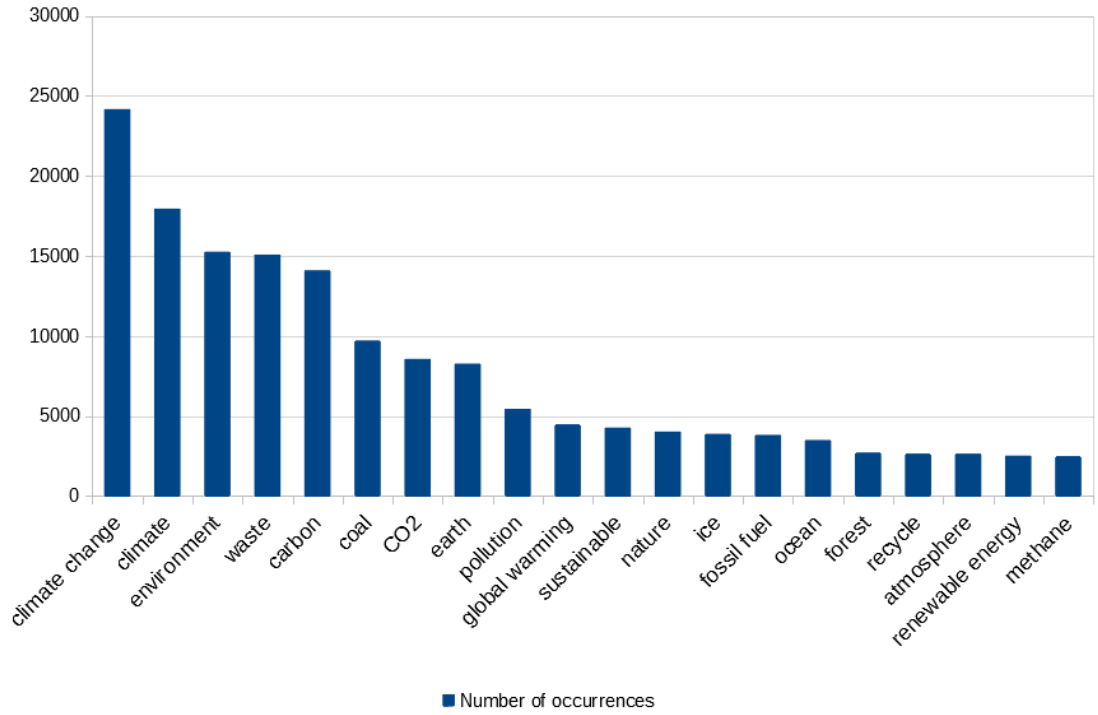

Figure 4.9 – Top 20 keywords that occur in the filtered Reddit data.

Furthermore, we found that the Reddit dataset contained fewer artifacts commonly found in Twitter data, such as links, emojis, user mentions, and hashtags. Figure 4.10 shows the number of occurrences of each artefact in Twitter and Reddit for a randomly selected set of 5,000 texts.

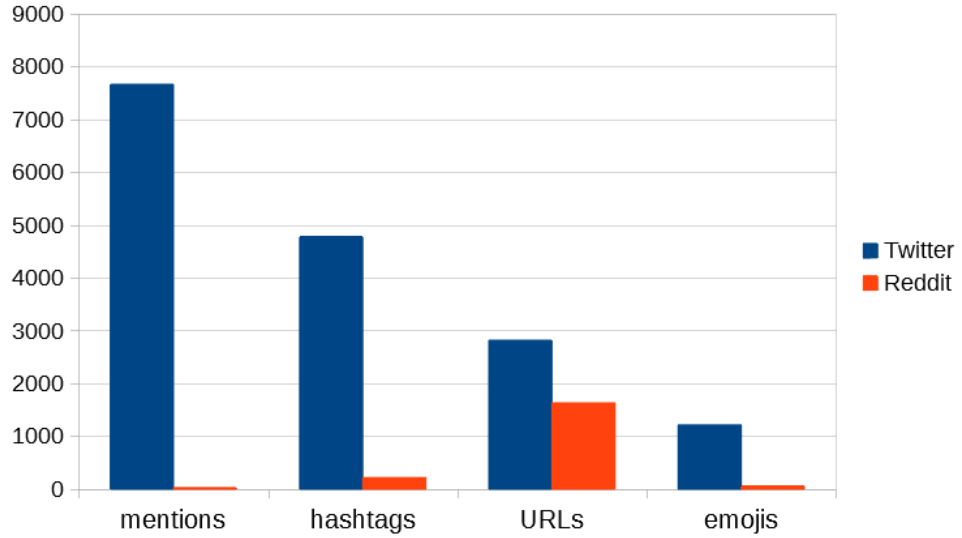

Figure 4.10 – Number of artifacts per 5,000 texts in Twitter and Reddit.

4.4 Results: corpora description

While studying the two social networks and performing data extraction, we have noticed certain similarities and differences between them. Regarding similarities, we can point out the availability of public data and acces to APIs. On the other hand, Twitter and Reddit are organized in different ways and, as a result, they offer different textual entities for data extraction. Communication on Twitter is based on tweets, while Reddit is structured around submissions and comments. These architechtural differences made us adopt two distinct approaches for building the respective corpora. For Twitter, we used a keyword-based approach, while for Reddit, we opted for a thematic communities-based approach.

An overview of the final corpora is presented in Table 4.6. It can be seen that the average length of a Reddit post is 46 tokens, whereas a tweet averages at 19 tokens. This is consistent with the fact that Twitter is a platform for SMS-like conversations, while Reddit functions as a forum-like social network. Further differences and similarities will be discussed in Chapter 6, where the corpora data will be used for a qualitative analysis of *carbon* in ordinary discourse.

The files containing Reddit post IDs and tweet IDs can be accessed through a public GitHub repository.[24] To access the actual content associated with these IDs, it is necessary to obtain developer access to the respective APIs, which can

[24]https://github.com/TG-PhD/LEGCOD_corpus_data.git

	Twitter	*Reddit*
Data type	tweets	submissions comments
Timeframe	02.09.2020 - 21.09.2020	17.09.2004 - 14.05.2021
Raw data	4,709,014 tweets	13,939 submissions 430,540 comments
Processed data	2,000,410 tweets	902 submissions 91,139 comments
Number of tokens	37,532,283	4,290,677
Average text length	19 tokens	46 tokens

Table 4.6 – General profile of two corpora collected from Twitter and Reddit.

be gained through the folowing links: 1) Twitter's API: https://developer.twitter.com/en/docs/twitter-api; 2) Reddit's API: https://www.reddit.com/dev/api/.

4.5 Chapter synthesis

In this chapter, we have addressed two main points. In the first part, we explained the challenging two-step process of keyword extraction. To avoid using the notion of Term and emphasize that we are dealing with abstract units with no clear terminological status, we introduced the concept of *extraction keyword* in the context of NLP. Additionally, we described the procedures of keyword extraction V1 and keyword extraction V2, highlighting the differences between two approaches. The final keyword list, comprising 290 items, is presented in Appendix A.

In the second part, we gave an overview of the two selected social networks for data extraction. We emphasized the similarities and differences between these platforms. Due to the difference in the way these platforms structure the information, we adapted the extraction setup accordingly for each social network. Furthermore, we discussed the ethical and legal considerations of using social media data for academic research. Both Twitter and Reddit are becoming more restrictive in terms of their APIs for numerous reasons. We outlined the procedure for filtering raw data from both networks which resulted in 2,000,410 filtered tweets and 92,041 Reddit sumbissions and comments that form a sample of online environmental discussions.

In Chapter 6, we exploit our social network corpora to create a "*carbon* subcorpus," which will be analyzed to study the usage of *carbon* and related expressions in ordinary discourse.

Chapter 5

Carbon in environmental agenda and *carbon* in the lexicon

Summary

Whereas the 20th century was the century of 'the gene' whose meaning has been studied by many social scientists [...], the 21st century will be the century of 'carbon' whose meaning needs to be studied, preferably before we enter the era of 'a post-carbon society'.

Nerlich, Koteyko and Brown (2010)

In this chapter, we embark on an in-depth linguistic profiling of *carbon* as a linguistic entity, with an additional focus on the sociological aspects related to the corresponding concept within the environmental domain. *Low-carbon* technological solutions occupy the experts' minds, as countries around the world strive towards *decarbonized* economy and society. As a result, *carbon* has become an emblematic environmental buzzword which remains ever-present in current news. We analyze what the concept of carbon represents in the environmental discourse. Furthermore, we conduct a linguistic analysis of *carbon* as a linguistic entity considering its usage by experts. Finally, we provide a formal lexicographic description of the various senses associated with *carbon*, including popular environmental idioms, such as ⌜*carbon dioxide*⌝, ⌜*carbon cycle*⌝, ⌜*carbon footprint*⌝, and ⌜*carbon capture and storage*⌝.

5.1 Carbon in chemistry

Scientists argue about the time when carbon atoms first appeared, with some stating it happened during the Bing Bang (Hazen 2019), while others argue for a later timeframe (NASA 2016). In prehistoric period, our ancestors encountered carbon in the form of charcoal and soot. However, they were unaware of these substances being made of carbon or their chemical properties. It was not until the 18th century that a series of experiments conducted by scientists revealed that diamond, graphite, soot, and charcoal[1] are forms of the chemical element carbon (C). While in chemical terminology it is criticizable to refer to chemical elements based on physical characteristics since those characteristics pertain to atoms and not elements themselves (Mikhel 2022: 130), it is a commonly used practice, and we will adopt it for the purpose of this discussion. However, we provide a detailed discussion of this terminological nuance in 5.5.2.

As asserted in Trifonov *et al.* (1982), the recognition of carbon as a chemical element dates back to 1789 when French chemist Antoine Lavoisier included carbon in his work *The Table of Simple Bodies.* Lavoisier's experiments with the combustion of diamond and charcoal allowed him to conclude that they were dif-

[1]Diamond and graphite are composed of pure carbon, whereas charcoal and soot are nearly pure forms.

ferent forms of the same element: carbon. Other chemists, including Smithson Tennant and Louis-Bernard Guyton de Morveau, confirmed his findings. In the book *Methods of Chemical Nomenclature* published in 1787, the authors used the name *carboneum* to refer to carbon (Trifonov *et al.* 1982). The name *carbon*, which originates from Latin *carbo* 'charcoal', was proposed in 1824.

Carbon is known as the *king of the elements* and *element of life* due to its omnipresence, abundance, and unique chemical properties. It is one of the most abundant chemical elements in the universe by mass and a key element of Earth's composition. Furthermore, carbon ranks as the second most common chemical element in the human body (accounting for 18,5% of body composition), surpassed by oxygen (Razeghi 2019). Consider how Hazen (2019) describes the omnipresence of carbon from the point of view of an ordinary person:

> Carbon is everywhere: in the paper of this book, the ink on its pages, and the glue that binds it; in the soles and leather of your shoes, the synthetic fibers and colorful dyes of your clothes, and the Teflon zippers and Velcro strips that fasten them; in every bite of food you eat, in beer and booze, in fizzy water and sparkling wine; in the carpets on your floors, the paint on your walls, and the tiles on your ceilings; in fuels from natural gas to gasoline to candle wax; in sturdy wood and polished marble; in every adhesive and every lubricant; in the lead of pencils and the diamond of rings; in aspirin and nicotine, codeine and caffeine, and every other drug you've ever taken; in every plastic, from grocery bags to bicycle helmets, cheap furniture to designer sunglasses. From your first baby clothes to your silk-lined coffin, carbon atoms surround you.

Carbon atoms possess the unique ability to form various bonds with one another as well as with atoms of other elements, resulting in a limitless number of compounds of different structures and chemical properties. There are around ten million known carbon compounds, with new ones being continuously synthesized by chemists. Together with hydrogen, nitrogen, oxygen, and other elements, carbon forms the most essential compounds for living things, such as carbohydrates, lipids, proteins and nucleic acids. In living organisms, these compounds form cells, store energy, and make up muscles, among other vital functions. These compounds are called organic and their chemical formulas contain both carbon

and hydrogen. Carbon is also present in inorganic compounds which do not contain carbon-to-hydrogen bond. Carbon monoxide and carbon dioxide are the most common examples of inorganic carbon-based compounds.

Industrial applications of pure carbon and carbon compounds are extensive. Diamonds find use in construction, mining, jewelry and oil-drilling, while graphite is used in production of refractory materials, as a machine lubricant and electrical conductor. Carbon dioxide is employed in food, agricultural, chemical, petrolium and health industries. Coal, oil and natural gas are utilized for heat and electricity production. Carbon fibers find applications in aerospace, medical and musical industries, to name just a few examples.

5.2 Carbon and the environment

A substantial portion of current environmental discussions revolves around carbon. Scientists are urging countries to reduce their *carbon emissions*, while individuals are being encouraged to control their personal *carbon footprint*. There is a growing trend in society as a whole to strive for a *low-carbon* and *decarbonized* future. Consider some recent news headlines from the British popular science magazine *NewScientist*:

- *Forests bioengineered to capture more **carbon** will be planted in the US.*
- *The race is on to tackle climate change by pulling **carbon** from the air.*
- *Most schemes to capture and reuse **carbon** actually increase emissions.*

In press, carbon is often depicted as something that should be *captured*, *reused*, and *pulled from the air*. In the environmental discourse, *carbon* is commonly (but not exclusively) used as a shorthand expression to refer to carbon dioxide (a detailed linguistic discussion on this topic will be provided in 5.4). For example, the ubiquitous expression *carbon emissions* usually refers to *carbon dioxide emissions*, or similar carbon-related expressions such as *carbon intensity* that implies *carbon dioxide intensity* (1).

(1) India has announced a target to reduce its **carbon intensity** (the amount of **carbon dioxide** released per unit of GDP) by 20-25% by 2020 from 2005 lev-

els. [OECD, *Transition to a low-carbon economy*, 2010, https://www.oecd-ilibrary.org/content/publication/9789264090231-en, p. 38]

The greenhouse gas carbon dioxide is arguably considered by scientists as the main culprit of rising atmospheric temperatures which leads to climate change. In the context of rising temperatures, scientists divide the history of humanity into two distinct periods: pre-industrial times (prior to 1900)[2] and the period from the early twentieth century until the present. Scientists claim that the long-term consequences of the industrial revolution have significantly changed the Earth's climate. Initiated in 1760, the industrial revolution spurred various human activities such as fossil fuels consumption, deforestation, and land use, among others. These activities are held accountable for the excessive amounts of carbon dioxide in the Earth's atmosphere, disrupting the carbon cycle and giving rise to a number of environmental issues. Specifically, Hazen (2019) points out the following "indisputable" facts about carbon's role in climate change:

1. Carbon dioxide and methane, which both contain the element carbon in their chemical composition, are powerful greenhouse gases. These gases have the ability to capture the Sun's radiation, leading to a decrease in the amount of energy that is radiated back into space. When there are higher levels of carbon dioxide and methane in the atmosphere, a greater amount of solar energy becomes trapped contributing to the greenhouse effect.
2. The levels of carbon dioxide and methane in Earth's atmosphere are increasing.
3. Almost all changes in atmospheric composition are being driven by human activities, particularly the combustion of billions of tons of carbon-rich fuels on an annual basis.
4. For over a century, Earth's surface temperature has been rising. Records dating back to 1880 indicate that the twelve warmest years have occurred within the past two decades.

Carbon has taken its firm place in environmental, scientific, economic, and

[2]For example, in IPCC (2018), the authors use *pre-industrial* to refer to the period of 1850-1900.

political agenda as we live in the era of *global carbon crisis* (Busch and Shrivastava 2011). Carbon's growing significance has given rise to such notions as *carbon colonialism* (Bachram 2004; Paterson and Stripple 2012), *carbocracy*, i.e., carbon technocracy (Lohmann 2001) and *carbonomics*, i.e., carbon economics (Busch and Shrivastava 2011). From a historical perspective, Levy and Spicer (2013) proposed to divide the trajectory of carbon history into two phases: the carbon wars of the 1990s (conflict within the energy sector and the rise of environmental issues) and the carbon consensus of the early 2000s (decarbonization of the economy). At the beginning of the twenty-first century, it has become clear that humanity is heavily dependent on carbon-based energy such as coal, oil, and natural gas. For example, China's economic success can be attributed to its abundant coal reserves and the availability of affordable hydrocarbons which have fueled its manufacturing, construction, and transportation industries, turning China into a manufacturing "hub." However, this reliance has made China one of the largest emitters of CO_2. As of 2021, the largest emitters of CO_2 were China, the United States, the European Union, India, Russia, and Japan, collectively accounting for 67.8% of global fossil CO_2 emissions (European Commission 2022).

Despite discouraging forecasts about the rise of Earth's temperature caused by carbon emissions, scientists still have hope that it is possible to repair the damage:

> Having the right policies, infrastructure and technology in place to enable changes to our lifestyles and behaviour can result in a 40-70% reduction in greenhouse gas emissions by 2050. (IPCC 2022)

Within the context of environmental mitigation and the urgent need to reduce global greenhouse gas emissions, the concepts of *low-carbon future* and *decarbonization* have emerged. These concepts refer to significant reductions in greenhouse gas emissions at the international, national, corporate, and individual levels. There are various examples of low-carbon technologies aimed at reducing greenhouse gas emissions. These technologies include more efficient use of natural resources, such as using heat pumps, reducing natural resource consumption, and exploring ways to utilize emissions, such as capturing and utilizing carbon dioxide or producing

goods using renewable resources.[3] One well-known low-carbon technology is *carbon dioxide capture and storage* technology (or *carbon capture and storage* technology, abbreviated as CCS). CCS refers to the process of separating CO_2 from industrial and energy-related sources, followed by its transportation to a storage location where it is securely isolated from the atmosphere for a long period of time (IPCC 2005). The significance of CCS for environmental mitigation has been compared to the moon landing (K. Buhr and Hansson 2011), highlighting its importance in the fight against climate change.

At the national level, numerous governments have made commitments to reduce emissions as prescribed in international agreements such as the Montreal Protocol (1987), the United Nations Framework Convention on Climate Change (1992), the Kyoto Protocol (1997), and the Paris Agreement (2015), among others. The European Union has developed the environmental strategy called *Climate Law*, which legally binds the member states to achieve climate neutrality, implying drastic reductions in greenhouse gas emissions, by 2050. At the corporate level, companies keep issuing announcements of their decarbonization goals, including such commercial giants as Google, Unilever, Volkswagen Group, and more.

5.3 Public engagement and carbon literacy

In section 2.3.5, we discussed public engagment and literacy in relation to environmental topic from a broad perspective. Let us now consider the public understanding of environmental issues associated with excessive greenhouse gas emissions, often referred to by the generic expression *carbon emissions*. Our personal lifestyle and choices have a significant environmental impact known as our *carbon footprint*. Approximately a quarter of global emissions originate from personal and household carbon footprints: households in Europe are responsible for 25% of greenhouse gas emissions (Jakucionyte-Skodiene, Krikstolaitis and Liobikiene 2022), and for 25% in the United States (Goldstein, Gounaridis and Newell 2020). The general public is actively engaged in carbon management to reduce these numbers. However, according to Ivanova and R. Wood (2020), only 5% of EU

[3] https://www.weforum.org/agenda/2022/02/what-are-low-carbon-emitting-technologies-an-expert-explains/

households are currently living within the European climate goals.

In the paradigm of a low-carbon society, individuals are expected to be *low-carbon consumers*, *low-carbon employees*, and *low-carbon citizens* (Whitmarsh, Seyfang and O'Neill 2011). Adhering to these principles may require individuals to reconsider their lifestyle – for instance, giving up taking flights and eating meat – or making small changes towards a low-carbon lifestyle, such as prioritizing local and organic food and limiting household energy consumption. Ivanova, Barrett, *et al.* (2020) analyzed 53 studies proposing ways to reduce personal carbon footprint and identified 61 most effective options for reducing greenhouse gas emissions. Figure 5.1 presents the top ten consumption options from different areas, including transportation (points 1, 2, 3, and 5), housing (points 4, 6, 8, and 10), and food (points 7 and 9) which, when combined, could significantly reduce excessive emissions.

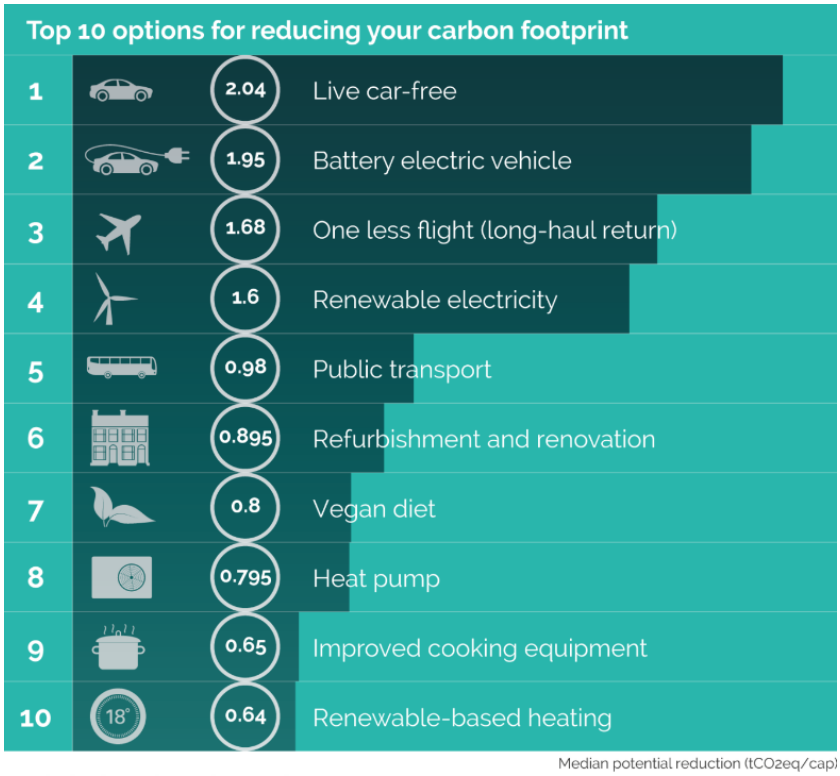

Figure 5.1 – Top ten options for reducing individuals' carbon footprint.[4]

Governments and business play a crucial role in proposing carbon reduction initiatives and mobilizing individuals to adopt these measures. However, these

[4]Source: https://environment.leeds.ac.uk/faculty/news/article/5471/global-study-uncovers-best-ways-to-change-consumption-to-cut-carbon-footprint

initiatives aimed at engaging the public do not always align with the level of public carbon literacy. The notion of *carbon literacy* was developed in the early 2000s by health professionals who recognized the opportunity to combine low-carbon lifestyle with healthy choices (Dósa and Russ 2020). Currently, carbon literacy is defined by the *Carbon Literacy Project* as the "awareness of the carbon costs and impacts of everyday activities and the ability and motivation to reduce emissions, on an individual, community and organisational basis."[5] Furthermore, public perception of the concept of *carbon* in carbon-related environmental issues remains an important question, as a deeper understanding of public's knowledge in this area enables policymakers to prepare well-informed policies and facilitates the implementation of low-carbon initiatives (Twyman, T. A. Smith and Arnall 2015).

Previous research has determined several factors that pose challenges to public carbon literacy. Specifically, these factors include a poor knowledge and understanding of terminology, emerging low-carbon technologies, and the underlying chemistry and physics of environmental issues. For example, a study by Gadema and Oglethorpe (2011) showed that 89% of respondents ($n = 428$) found carbon labeling[6] of goods to be confusing. In another study conducted by Hartikainen *et al.* (2014), it was discovered that 90% of respondents ($n = 1010$) were unable to provide a definition of *product carbon footprint*. As asserted in Hedin, Grönborg and Johansson (2022), public knowledge about greenhouse gas emissions associated with food production and transportation is low. To address this, the authors proposed the use of a digital grocery list for mobile phones, which resulted in improved food carbon literacy among the users as they became more knowledgeable at estimating carbon footprint of food products utilizing the application.

Grinstein *et al.* (2018) conducted a study on *carbon numeracy*, which refers to people's quantitative understanding of their carbon footprint without relying on detailed calculations, and how it influences their consumer choices. The study showed that individuals struggle to estimate the carbon footprint associated with

[5]https://carbonliteracy.com/what-on-earth-is-carbon-literacy/

[6]Carbon labeling emerged as a form of eco-certification in the early 2000s. Nowadays, it concerns the labeling of various goods, airlines and train services, car rentals, and hotels to provide information about their environmental impact, i.e., carbon footprint.

driving a motor vehicle. This kind of innumeracy was found to be higher compared to other metrics, such as counting calories and distance. The author concluded that ensuring consistency in carbon information and carbon footprint calculations, enhancing knowledge about carbon footprint and its impact, and utilizing alternative metrics for carbon footprint assessment (such as carbon points or energy points) can all contribute to carbon numeracy.

Furthermore, Coulter *et al.* (2007) highlighted the confusion of the general public when confronted with emission-related terminology, such as *carbon*, CO_2, *carbon dioxide*, and *carbon equivalent*. In Table 5.1, which is borrowed from Coulter *et al.* (2007), you can find a summary of public attitudes towards environmental terminology, including terminology related to carbon emissions.

	Positive aspects	**Negative aspects**
Global warming	Familiar	Old fashioned Misleading Abstract
Climate change	Familiar Current	Misleading Abstract
Carbon emissions	Able to visualize	Unfamiliar Misleading
Carbon footprint	Conveys personal impact once understood	Unfamiliar
Carbon neutrality	Thought provoking	Confusing Technical
Carbon offsetting	Thought provoking	Confusing Technical

Table 5.1 – Summary of public attidues towards environmental terminology (Coulter *et al.* 2007).

Public understanding of low-carbon technologies is of particular importance to scientists and educators as the implementation of these technologies often depend on public acceptance. Unfortunately, environmental initiatives aimed at reducing carbon emissions often lack public support. One notable example is the Australian carbon tax, which was abolished several years after its implementation due, in part, to public and industry opposition (Lopes, Antunes and Janda 2020). Similarly, France also had to eliminate its carbon tax. Furthermore, a number of carbon capture and storage projects, such as *Schwarze Pumpe* in Germany, *Greenville* in the USA, *Barendrecht* in the Netherlands, and *CarbonNet* in Australia, were

rejected by the public (Simonchuk and Romasheva 2021).

As asserted in Murray (2022), the lack of public acceptance is often a result of lack of knowledge on the subject. In the case of carbon capture and storage technologies, public disapproval may be attributed to a poor understanding of the chemical and physical properties of carbon dioxide (Wallquist, Visschers and Siegrist 2009). Additionally, a study by Ha-Duong, Nadaï and Campos (2009) revealed that only 6% of the respondents ($n = 1076$) were able to provide a correct definition of carbon capture and storage technology.

To address this knowledge gap, various educational initiatives have been established. For example, the *Carbon Literacy Project* offers courses for organizations, individuals and educators, aiming to improve carbon literacy. To date, over 59,000 citizens worldwide have completed their courses.

To sum up, the focus of the current environmental debate is on carbon emissions, which commonly stands for carbon dioxide emissions. Scientific findings concerning the Earth's rising temperatures emphasize the necessity for collaborative international, governmental, and public action. To raise public awareness about the issue of excessive greenhouse gas emissions, governmental and corporate initiatives aimed at reducing one's personal carbon footprint have become commonplace. However, research on public literacy shows that the general population lacks basic scientific knowledge on the environmental issues and finds *carbon*-related terminology unclear. In 5.4, we will discuss the underlying terminological factors that contribute to this kind of confusion among the public.

5.4 Linguistic analysis of *carbon*: origins of terminological confusion

By conducting an in-depth qualitative linguistic anlysis of *carbon* in scientific and educational texts, we have identified the following factors that contribute to the complication of *carbon*-related terminology and, as a result, confusion among the general public:

1. ambiguity of *carbon* and related idioms in specialized discourse;
2. interchangeable use of synonymic terminological variants;
3. definitional lacuna in lexicographic sources.

In the following sections, we will discuss each point individually and provide illustrative examples from scientific literature. The analysis of the public perception of *carbon* will be discussed in Chapter 6.

5.4.1 The ambiguity of *carbon*

In section 5.1, we introduced the chemisty-related sense of *carbon*, which refers to the chemical element. Within the environmental domain, the chemical element carbon is commonly referenced as:

- the main participant in the Earth's carbon cycle, which is being disrupted by human activity, as in (2);
- one of the constituents of carbon-containing greenhouse gases, such as carbon dioxide (CO_2) and methane (CH_4), as in (3).

(2) In the carbon cycle, the element **carbon** is constantly recycled between the Earth and its atmosphere, helping to sustain life on our planet.
[Web, https://www.solarschools.net/knowledge-bank/climate-change/carbon-cycle, 20/08/2023]

(3) Two of the greenhouse gases (carbon dioxide and methane) contain **carbon** atoms.
[Web, https://www.energy.gov/nepa/articles/ea-1162-final-environmental-assessment, 20/08/2023]

Additionally, as we mentioned previously, *carbon* is frequently used as an elliptical expression referring to the greenhouse gas carbon dioxide. Specifically, this usage proliferates in the idiom *carbon emissions*, which commonly (but not exclusively) denotes the emissions of carbon dioxide. It is difficult to pinpoint the exact moment when *carbon* first became a common shorthand for *carbon dioxide* in the environmental discourse. Regarding *carbon emissions*, the *Oxford English Dictionary* provides its first occurrence in 1921, but citations relevant to the environmental discourse date back to 1977:

(4) Table 1 shows the calculated U.S. energy consumption and world **carbon emissions** along the uncontrolled and controlled paths. [*OED*, online, 2023]

The logic behind this shorthand form can be explained from two perspectives:

- the abbreviated form results from omitting "dioxide" in *carbon dioxide*;
- the form *carbon* is a result of metonymy, because carbon dioxide is a gas made up of two elements: carbon and oxygen.

The shorthand form *carbon* is commonly used across various registers of environmental discourse, ranging from scientific and mass media discourse to public discourse. However, scientific community advocates for more precise terminology. This implies that using *carbon dioxide* is preferred over *carbon*, and *carbon dioxide emissions* is preferred over *carbon emissions*. In certain cases, the experts specify the adopted terminology to enhance the precision of their technical writing, as in the following example (5).

(5) The terms "**carbon**," "**carbon dioxide**," and "**CO2**" are used interchangeably throughout this paper, but the design of the policy modeled in this paper is based on CO2 emissions. The relationship between carbon and carbon dioxide is based on the atomic weights of the components such that a policy that placed a price on carbon alone would be approximately 12/44ths the price on carbon dioxide emissions.
[Perese Kevin, *Input-output model analysis: pricing carbon dioxide emissions*, 2010, Congressional Budget Office]

Despite the recommendation to avoid the shorthand expression *carbon*, the corpus search with the *Google Ngram Viewer*[7] indicates that the form *carbon emissions* has been prevalent since the early 2010s (see Figure 5.2).

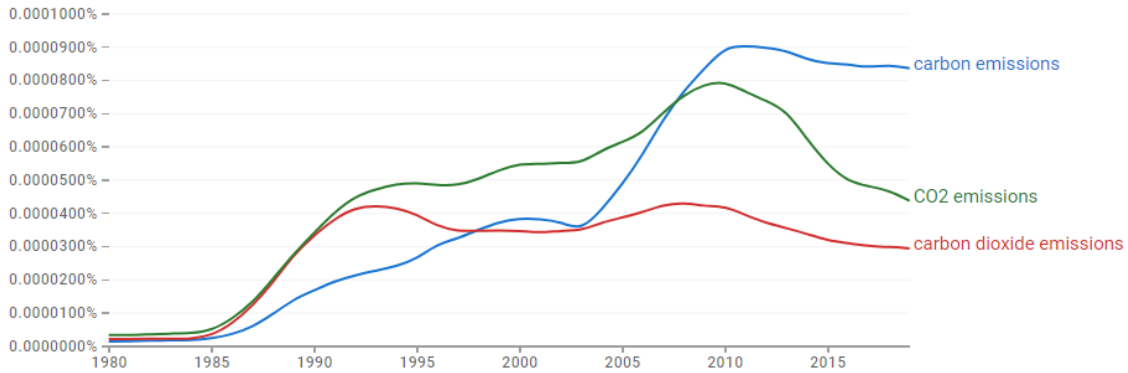

Figure 5.2 – The occurrence of *carbon emissions*, *carbon dioxide emissions* and *CO_2 emissions* in *Google Books Ngram Viewer* over the period of 1980-2019.

The usage of the shorthand form *carbon* 'carbon dioxide' and *carbon* 'chemical

[7] https://books.google.com/ngrams/

element' within a single context can be misleading. In (6), the intended meaning is clear, with the first instance of *carbon* denoting a greenhouse gas carbon dioxide, and the second instance referring to the chemical element. However, in many other contexts, lack of clear indications regarding the meaning of *carbon* might create terminological ambiguity.

(6) At the same time, changes in the Earth's climate system are altering forests in dramatic ways, which can also have consequences for the emission of **carbon** and other greenhouse gases. [...] **Carbon** is one of the most important elements found on Earth. [Web, https://www.fs.usda.gov/ccrc/topics/global-carbon, 16/08/2023]

As previously mentioned, carbon dioxide is not the only greenhouse gas that contains carbon. The greenhouse gas methane, which has recently gained environmental prominence due to its potency, also contains the chemical element carbon. In some contexts, *carbon* is used as a shorthand expression to denote both carbon dioxide and methane (7), further complicating *carbon*-related terminology.

(7) Current permafrost areas are projected to be a net emitter of **carbon** (CO_2 and CH_4). [IPCC, *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 2014, p. 74]

Furthermore, the shorthand form *carbon* is used to refer to the carbon-containing gas carbon monoxide (CO), which is not classified as a greenhouse gas, although it indirectly contributes to climate change (8).

(8) We describe a new rural network of four densely placed ($<$ 100 km apart), continuous atmospheric **carbon** (CO_2, CH_4, and CO) measurement sites in north-central Switzerland and analyze its suitability for regional-scale ($\sim$ 100–500 km) carbon flux studies.
[Oney Brian *et al.*, *The CarboCount CH sites: characterization of a dense greenhouse gas observation network*, 2015, Atmospheric Chemistry and Physics, Vol. 15, p. 11147]

Thus far, we have identified two distinct senses of *carbon* in environmental discourse: *carbon* as a chemical element and *carbon* as a gas, specifically referring to greenhouse gases such as carbon dioxide, methane, and carbon monoxide. However, when used in idiomatic expressions, *carbon* acquires additional semantic nuances.

Carbon-related phraseology. The prevalence of environmental issues related to excessive carbon dioxide emissions has significantly contributed to the widespread use of *carbon* idioms, also referred to as "carbon compounds" in linguistic research (Koteyko 2012). The following citation shows how extensive and versatile *carbon*-related phraseology is and how it can encompass a wide range of environmental concepts:

> To be *carbon-neutral*, you can use a *carbon calculator* to estimate your household's *carbon footprint*. Then you can seek to reduce your own *carbon emissions*, or you can purchase *carbon offsets* or *carbon credits*. Countries can institute *carbon taxes*, while eco-conscious companies can engage in *carbon trading* on the *carbon market*. And maybe someday, if we're all *low-carbon* or even *zero-carbon*, we can live in a *post-carbon* world. (Zimmer 2007)

While the idioms mentioned in this citation are relevant primarily to environmental mitigation, it is important to note that *carbon*-related idioms can refer to a broader range of scientific concepts (see Table 5.2). Additionally, there are also popular lay idioms that continue to emerge, often used for public engagement and marketing purposes, such as *carbon diet*, *carbon detox*, and *carbon foodprint*.

black carbon	*blue carbon*
carbon budget	*carbon negative*
carbon capture	*carbon price*
carbon capture and storage	*carbon removal*
carbon cost	*carbon sequestration*
carbon cycle	*carbon sink*
carbon dioxide	*carbon trade*
carbon economy	*embodied carbon*
carbon free	*net zero carbon*
carbon intensive	*operational carbon*
carbon mitigation	

Table 5.2 – Examples of *carbon* idioms.

In Table 5.2, the idioms possess different phraseological status. For example,

the majority of them are clearly *weak idioms*.[8] This means that the meaning of these idioms contains the meanings of their respective components as well as additional semantic material. For instance, the meaning of the idiom *carbon cycle* contains the meaning of *carbon* 'atoms of element C' and *cycle* 'repetitive motion.' In contrast, the idiom *blue carbon* that refers to 'carbon sequestered in the ocean' (Zinke 2020) is a *semi-idiom*. Its meaning does not entail any of the existing senses of *blue* as none of them refers to the ocean. Similarly, *blue* in *blue carbon* does not refer to the color of carbon itself.

Our semantic analysis of *carbon* as a component of idioms has revealed that it exhibits greater semantic versatility than the standalone lexical unit *carbon*. Specifically, within idiomatic expressions, *carbon* can refer to a wider range of greenhouse gases beyond just carbon dioxide and methane. Firstly, it can refer to specific greenhouse gases, such as carbon dioxide, methane, and nitrous oxide. For instance, in (9), *carbon footprint* includes the emissions of carbon dioxide, methane, and nitrous oxide. Similarly, the Australian Carbon Farming Initiative[9] scheme considers methane and nitrous oxide to be part of *carbon emissions*.

(9) A '**carbon footprint**' is an estimate of all the greenhouse gases (GHGs) associated with a process or product. It converts emissions of individual GHGs into a single carbon dioxide equivalent(CO_2eq) value using the global warming potential (GWP) of the individual gases over a 100 year period. Although there are numerous GHGs, the footprint, especially in a farming and food context, represents the total emissions of carbon dioxide (CO_2), methane (CH_4) and nitrous oxide(N_2O). [Lillywhite Robert, Collier Rosemary, *Why carbon footprinting (and carbon labelling) only tells half the story*, 2009, *Aspects of Applied Biology*, Vol.95, p. 73]

Secondly, *carbon* in idioms can refer to a group of greenhouse gases without specifying which ones precisely or even refer to all greenhouse gases collectively. In (10), *carbon* in *carbon constraints* is defined with "non-carbon-containing greenhouse gases" without further explanation.

(10) [...] our use of "carbon" in the term **carbon constraints** does not follow a narrow definition; it also covers other non-carbon-containing greenhouse gases as well as the direct negative effects of climate change. [Busch Timo, Shrivastava

[8]For further and more detailed typology of idioms, see Mel'čuk (2023).

[9]https://www.cleanenergyregulator.gov.au/Infohub/CFI/Carbon-Farming-Initiative

Paul, *The global carbon crisis : emerging carbon constraints and strategic management options*, 2011, Greenleaf Publishing Sheffield, UK, p. 13]

It is important to note that the official list of greenhouse gases issued by the Kyoto Protocol includes the following gases:

- carbon dioxide (CO_2);
- methane (CH_4);
- nitrous oxide (N_2O);
- hydrofluorocarbons (HFCs);
- perfluorocarbons (PFCs);
- sulphur hexafluoride (SF_6).

To our knowledge, none of the existing definitions of *carbon* in scientific and educational sources mention greenhouse gases that do not contain carbon in their chemical formulas, e.g., nitrous oxide and sulphur hexafluoride. However, as explained earlier, certain idioms do refer to such gases. Consequently, in such instances, the meaning of *carbon* cannot be inferred without explicit indications.

Furthemore, beyond the fluctuating meaning of *carbon* within idioms, the syntactic structure of certain lay idioms can be misleading, as they allow for multiple interpretations. For example, the British restaurant chain *Las Iguanas* offers *carbon-friendly cocktails* and pledges to plant trees for every bottle of spirit sold to offset carbon emissions.[10] The expression *carbon-friendly* in the idiom *carbon-friendly cocktails* follows the pattern of *environment-friendly*, which commonly refers to something that does not harm the environment. Without proper context, *carbon-friendly* may be interpreted as something which does not harm carbon itself.

In conclusion, our analysis has revealed that within environment-related idioms, *carbon* refers to multiple things, including the chemical element C, greenhouse gases, and carbon monoxide (see Figure 5.3). A notable distinction between *carbon* as a standalone lexical unit and *carbon* within idioms is that the latter can

[10]https://www.iguanas.co.uk/carbon-friendly

refer to a wider range of greenhouse gases, including those that do not contain carbon.

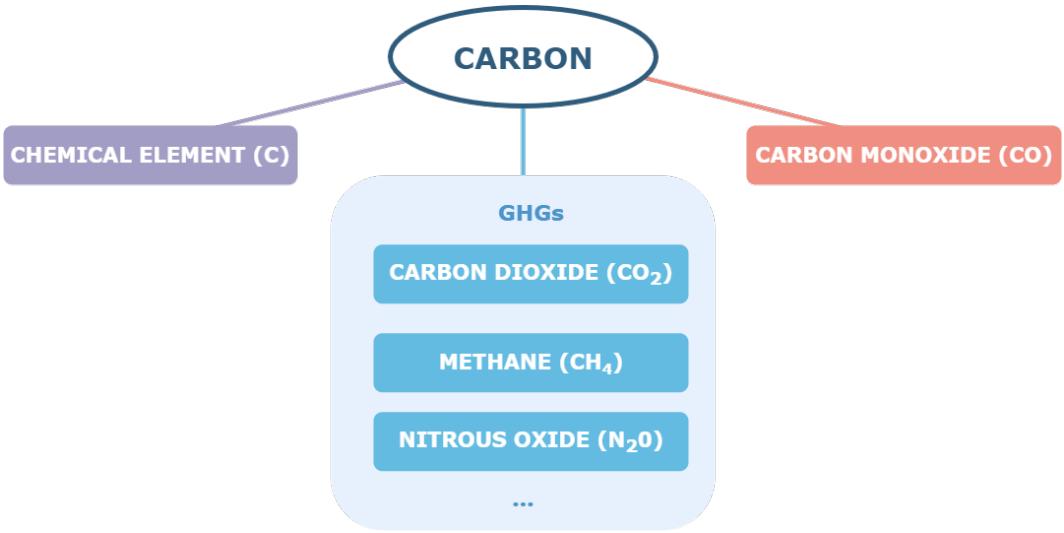

Figure 5.3 – The variation of senses of *carbon* within idioms.

Despite this semantic versatility, *carbon* idioms in educational sources are often narrowly defined with a specific focus on carbon dioxide. In Figure 5.4, the authors of the educational leaflet *Carbon Jargon Buster* provide a definition of *carbon emissions*, specifically referring to CO_2 and other greenhouse gases. This idiom is further embedded in the definition of *carbon footprint*, but the accompanying illustration portrays the greenhouse gas CO_2 only.

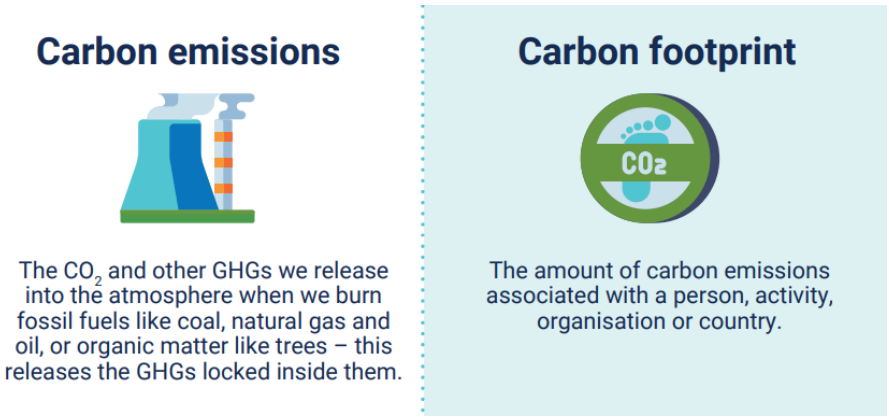

Figure 5.4 – Definitions of *carbon emissions* and *carbon footprint* in *Carbon Jargon Buster*.[11]

[11] https://www.josephash.co.uk/wp-content/uploads/2022/06/Joseph-Ash-Galvanizing-Carbon-Jargon-Buster-Poster.pdf

Similarly, in Figure 5.5, the idiom *carbon emissions* in the educational leaflet *Carbon terminology explained* is ambiguously defined with *carbon*, yet the corresponding illustration only features the greenhouse gas CO_2.

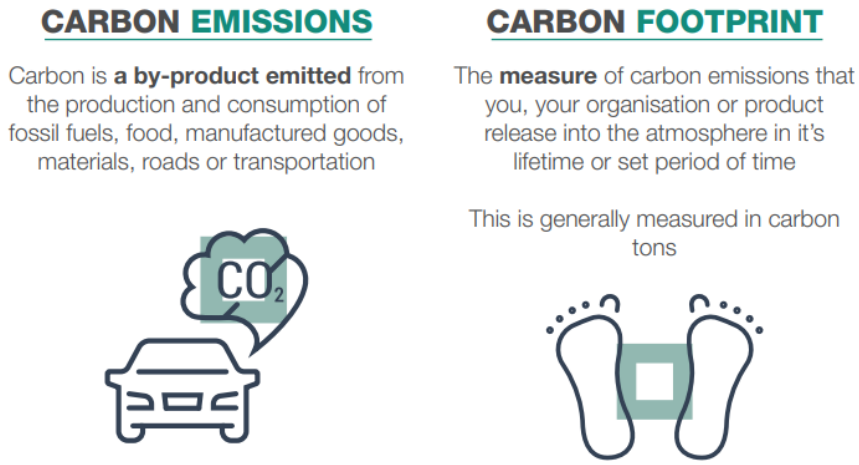

Figure 5.5 – Definitions of *carbon emissions* and *carbon footprint* in *Carbon terminology explained*.[12]

In Section 5.5, we focus on the lexicographic analysis of three specific idioms: ⌜*carbon cycle*⌝, ⌜*carbon footprint*⌝, and ⌜*carbon capture and storage*⌝. Our aim is to illustrate the nuances in the meaning and reference of *carbon* within each of these idiomatic expressions.

5.4.2 Terminological variation

In addition to the ambiguity of *carbon* in weak idioms, such idioms are prone to terminological variation. There are two types of variation pertinent to specialized lexicon: *denominative variation* and *conceptual variation* (Cabezas-García and León-Araúz 2023). *Denominative variation* in specialized language occurs when different designations are used to refer to a single concept. On the other hand, *conceptual variation* arises when a single designation is used to describe different concepts. The origins of terminological variation lie in the varying viewpoints of experts, the diversity of text genres used to address specialized topics, and the author's background and expressive intentions (Freixa and Fernández-Silva 2017).

[12]https://m-ar.co.uk/wp-content/uploads/2020/11/Carbon-Terminology-download.pdf

Terminological variation is particularly common in fast-developing and dynamic specialized fields where the normalization of terminological nomenclatures often lag behind the field advancements. Environmental science, for example, constantly introduces new Terms and concepts to account for emerging solutions to environmental issues. Additionally, environmental science exists at the intersection of scientific language, simplified mediation language, and general public language. This dynamic and versatile context contributes to the emergence of new designations for the same concept.

However, terminological variation hinders effective science communication. The lack of consistency and unsystematic representation of environmental terminology creates barriers to public engagement and participation in environmental mitigation efforts. It also hampers public knowledge and public awareness, leading to confusion and misunderstanding. The general public can become overwhelmed by the proliferation of environmental Terms, especially when they are used interchangeably and lack proper definitions.

The notion of domestication is particularly relevant in the context of terminological variation, as it involves various genres and levels of specialization, with speakers possessing varying degrees of expertise in the domain (Humbert-Droz and Picton 2022).

Denominative variation, firstly, originates from the use of the shorthand form *carbon* instead of the full form *carbon dioxide*. It is common for authors to interchangeably use *carbon emissions* and *carbon dioxide emissions* throughout the text when referring to the emissions of carbon dioxide. Similarly, the technology known as *carbon capture and storage* is synonymous with *carbon dioxide capture and storage*. However, substituting *carbon* with *carbon dioxide* does not apply to all cases. For example, the expression *carbon* ***dioxide*** *cycle* necessarily refers to the circulation of carbon dioxide, which is only a part of the broader *carbon cycle*.

Secondly, environmental discourse abounds with synonymous expressions that lack clear delineation of their meaning – for example:

- the concept of human detrimental impact on the environment can be designated by *anthropogenic footprint*, *environmental footprint*, *global footprint*,

carbon footprint, *carbon dioxide footprint*, *greenhouse gas footprint*, *footprint*, *emissions footprint*;

- the concept of corporate or personal financial responsibility for environmental damage is represented by *carbon tax*, *environmental tax*, *green tax*;
- the concept of something being environmentally beneficial or not harmful is designated by *sustainable*, *green*, *carbon neutral*, *carbon negative*, *clean*, *zero carbon*, *net zero*, *net zero carbon*, *climate positive*.

Such expressions allow for multiple interpretations, which is the reason why they can be exploited by companies seeking to remain vague about their environmental claims, a practice known as *greenwashing*. The Environmental Communication Guide developed by the company European Bioplastics recommends avoiding expressions such as *green*, *sustainable*, *environmentally friendly*, and *climate friendly* to reduce the ambiguity of environmental claims.[13]

While certain positive aspects of terminological variation can be identified, such as using *climate emergency* instead of *climate change* to create public response (Cabezas-García and León-Araúz 2023), it is important to create clear and consistent language across the channels and mediators to ensure effective communication between scientisits, policymakers, and the general public. Potential solutions to address variation in environmental terminology include international terminological unification and normalization, as well as the representation of lexical variants in lexicographic sources.

5.4.3 *Carbon* in current lexicographic sources

In 5.5, we present our lexicographic analysis focused on modeling the vocable CARBON in the English Lexical Network. However, it is important to first review how current lexicographic sources define *carbon* and provide an overview. For this, we conducted comparative analysis of general and specialized lexicographic sources related to the environmental domain, aiming to observe how *carbon* is treated in these sources. The list of the lexicographic sources used can be found in Table 5.3.

[13]https://www.european-bioplastics.org/news/publications/environmental-communication-guide/

Author / Publisher	**Title**	**Year**
Specialized resources		
EEA	*General Multilingual Environmental Thesaurus*	2021
EPA	*Collection of terminological vocabularies*	2021
LexiCon Research Group	*EcoLexicon*	2021
M.-C. L'Homme	*DicoEnviro*	2021
UNFCC	*Glossary of climate change acronyms and terms*	2021
IPCC	*Glossary to Special report Global Warming* $1.5°C$	2018
IPCC	*Glossary to Synthesis report Climate Change 2014*	2014
M. Allaby	*Oxford Dictionary of Environment and Conservation*	2013
M. Allaby	*Oxford Dictionary of Ecology*	2010
J. Daintith	*Oxford Dictionary of Chemistry*	2008
A. Porteus	*Dictionary of Enironmental Science and Technology*	2008
IPCC	*Glossary to Synthesis report Climate Change 2007*	2007
P.H. Collin	*Dictionary of Enironment and Ecology*	2004
McGraw-Hill	*Dictionary of Enironmental Science*	2003
T. Pankratz	*Environmental Engineering Dictionary and Directory*	2001
D. Kemp	*The Environment Dictionary*	1998
A. Gilpin	*Dictionary of Environment and Sustainable Development*	1996
Eurostat	*Environment Glossary*	N/A
General language resources		
Merriam-Webster	*Merriam-Webster Online Dictionary*	2021
Oxford University Press	*Lexico*	2021
Collins	*Collins Online English Dictionary*	2021
M. Rundell	*Macmillan Dictionary*	2021
Cambridge University Press	*Cambridge Dictionary*	2021
Pearson	*Longman Dictionary of Contemporary English*	2021
Houghton Mifflin Harcourt	*The American Heritage Dictionary*	2020

Table 5.3 – Lexicographic sources used for a comparative analysis of definitions of *carbon*.

Environmental organizations such as the IPCC and the UNFCC, that are responsible for some of these sources, are aimed at bulding environmental awareness. Therefore, we believe it is particularly important for them to maintain a method-

ological approach when providing definitions and deciding on the headword list.

We concluded that the overall treatment of *carbon* is not systematic and lacks adequacy in some cases. We identified two major shortcomings in the treatment of *carbon* within the analyzed lexicographic sources. First, the absence of *carbon* as a headword in the source's nomenclature was observed in certain specialized sources, such as the IPCC glossaries. Conversely, all the general dictionaries we consulted include *carbon* as a separate entry. Additionally, we found that the sources used generally focus on only one sense of *carbon*, despite its polysemous nature.

The primary issue lies in the fact that *carbon* is not included as an individual entry in the nomenclature of these sources, despite being present as part of certain idioms: *carbon budget, carbon cycle, carbon intensity, carbon neutrality* (as found in the *IPCC Glossary to Special report Global Warming* $1.5°C$), *carbon cycle, carbon sequestration* (as found in the *IPCC Glossary to Synthesis report Climate Change 2014*), *carbon market, carbon sequestration* (as found in the *UNFCC Glossary of climate change acronyms and terms*). We argue that it is inconsistent to have an entry for an idiom (such as *carbon cycle*) that contains a simple Term (such as *carbon*), without having a separate entry for that simple Term.

The following sources demonstrate a better practice of describing *carbon*:

- *Dictionary of Environment and Ecology* by P.H. Collin,
- *Environmental Engineering Dictionary and Directory* by T. Pankratz,
- *The Environment Dictionary* by D. Kemp,
- *Oxford Dictionary of Chemistry*,
- *Dictionary of Enironmental Science and Technology* by A. Porteus,
- EPA's *Student's Guide to Global Climate Change*,
- *General Multilingual Environmental Thesaurus*,
- *Oxford Dictionary of Ecology* and *Oxford Dictionary of Environment and Conservation* by M. Allaby,
- *DicoEnviro*,
- *EcoLexicon*.

These dictionaries acknowledge the importance of including *carbon* as a separate entry in their headword lists. Nevertheless, it is worth noting that these

sources only address one specific sense of *carbon*, despite the fact that it can refer to multiple senses pertinent to the environmental domain. We suppose that it can be explained by the fact that polysemy is usually ignored in terminological sources. However, we argue that there are several senses of *carbon* that may be included in a specialized lexicographic source on the environment. In contrast, all the general language dictionaries analyzed include *carbon* as an individual entry and provide at least a few senses associated with it.

Journalists partly address the necessity of clarifying the senses attributed to *carbon* and related idioms to the general public. We found a number of news articles dedicated to "debunking the carbon jargon" (see Figure 5.6). In such articles, the authors define commonplace environmental Terms such as *zero emissions*, *carbon negative*, *greenwashing*, *embodied carbon*, *carbon pricing*, etc. They put an emphasis on the complexity of existing terminology and address specific points of confusion.

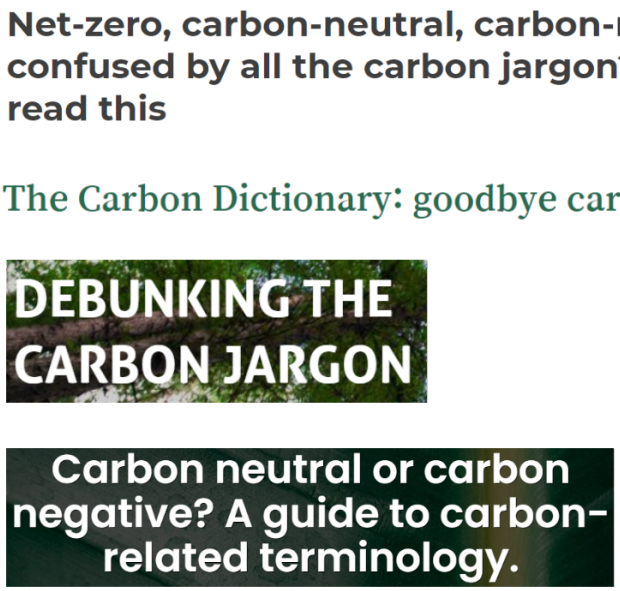

Figure 5.6 – The headers of news articles aimed at clarifying *carbon* terminology.[14]

[14]The respective news articles: 1) https://theconversation.com/net-zero-carbon-neutral-carbon-negative-confused-by-all-the-carbon-jargon-then-read-this-151382;
2) https://www.pawprint.eco/eco-blog/carbon-dictionary-goodbye-carbon-jargon;
3) https://www.proximagroup.com/debunking-the-carbon-jargon/;
4) https://www.lowcarbonmaterials.com/blog/carbon-neutral-carbon-negative-guide-to-carbon-related-terminology.

5.4.4 Overview of the polysemy of the vocable CARBON

While accumulating knowledge about the use of *carbon* in specialized texts and ordinary discourse (we provide the detailed study in Chapter 6), we started drafting a list of all senses associated with *carbon*. Furthermore, we used this information to perform the lexicographic work in the English Lexical Network. As a result, we propose the following polysemy of the vocable CARBON:

- **spec** CARBON **I.1** 'element **III.3a** with ⌜atomic number⌝ 6'
  [*Organic molecules contain carbon atoms.*]
- **(spec critic)** CARBON **I.2** 'substance **I.1a** that is the materialization of carbon **I.1**' [*Carbon is a black solid.*]
- **(spec critic)** CARBON **II.1** 'gas containing carbon **I.1** atoms'
  [*Several local facilities pledged to cut their carbon.*]
- **quasi-spec** CARBON **II.2** 'symbolic substance as if it were carbon **II.1**' [*Rich people use too much carbon.*]
- CARBON **III.1** '⌜carbon paper⌝' [*In the past, the writers typed the text on carbons.*]
- CARBON **III.2** '⌜carbon copy⌝' [*Mary keeps carbons of her letters.*]

The vocable CARBON is a terminological vocable, which implies that its polysemy is structured around the specialized basic lexical unit[15] – CARBON **I.1**.

In 5.5, we provide a detailed description of each sense. While the lexemes CARBON **I.1**, CARBON **I.2**, CARBON **II.1**, and CARBON **II.2** are relevant to the environmental domain, the lexemes CARBON **III.1** and CARBON **III.2** fall outside the scope of the environmental topic, which is why we do not discuss them further in the text.

All the senses within the polysemous vocable CARBON are connected with copolysemy relations (see Figure 5.7). *Copolysemy relations*, such as *Extension* and *Metaphor*, model the internal polysemy structure of the vocable by connecting the lexical units grouped within the given vocable (Polguère 2018).

[15] According to the definition in Mel'čuk and Polguère (2018: 420), "the basic lexical unit of the vocable is the lexical unit that semantically controls the polysemy structure of that vocable."

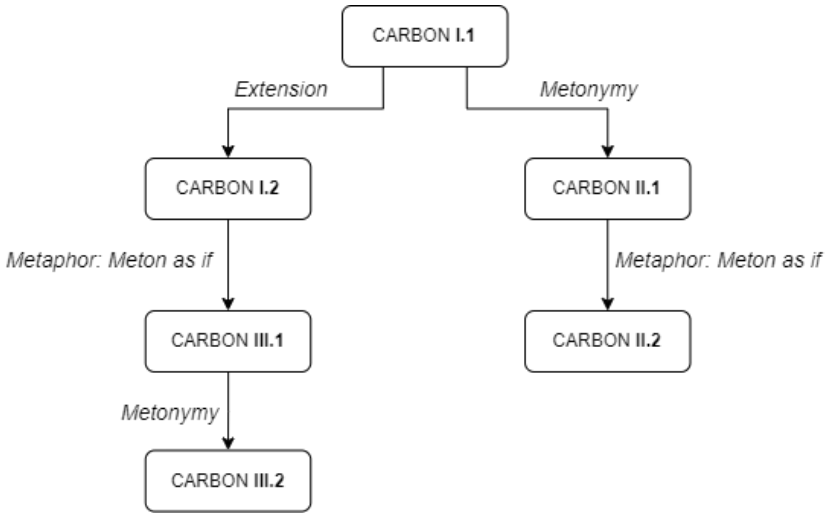

Figure 5.7 – Polysemy structure of the vocable CARBON.

5.5 Lexicographic description of the vocable CARBON and related idioms in the English Lexical Network

5.5.1 Methodology

The comparative analysis of how lexicographic sources handle *carbon*, as discussed in 5.4.3, revealed various limitations. Certain sources do not provide a separate entry for *carbon*, while others provide only a single sense. After highlighting the weaknesses of these approaches, we can now proceed with our suggested systematic treatment of *carbon*.

Specifically, our approach involves employing the principles of the Explanatory and Combinatorial Lexicology (see 1.3 and 3.1.3) for the lexicographic treatment of lexical units. The lexicographic aspect of work was performed in the English Lexical Network (en-LN), which is a Lexical System designed to structure the English lexicon, with a focus on its relational properties (Polguère 2014). To access the English Lexical Network, we used the lexicographic software *Dicet*, which enables us to perform lexicographic tasks and browse the data in the network (Gader, Lux-Pogodalla and Polguère 2012).

Considering the linguistic nuances of *carbon* and *carbon*-related expressions in specialized and ordinary discourse,[16] we have selected the following vocables for

[16]Certain aspects of the lexicographic analysis rely on the findings obtained from the study of ordinary discourse that will be presented in the following Chapter 6.

the lexicographic treatment of their corresponding lexemes:

- CARBON;
- ⌜CARBON DIOXIDE⌝;
- ⌜CARBON CYCLE⌝;
- ⌜CARBON FOOTPRINT⌝;
- ⌜CARBON CAPTURE AND STORAGE⌝.

The lexicographic articles that follow are intended to serve as suggestions and foundations for the further development of *carbon*-related terminology. The proposed definitions were formulated on the basis of the knowledge that we accumulated throughout the entire research process. It is important to note that we are not experts in the environmental domain and had to study the fundamentals of environmental science and specificities of environmental discourse. Additionally, a part of the lexicographic work was done collaboratively with experts in lexicology and green chemistry.

An entry for a lexeme in the English Lexical Network comprises the following data:

- grammatical information: part of speech, gender, phraseological status, etc.;
- usage notes: **spec**, **(spec)**, **(quasi-spec)**, etc.;
- propositional form which accounts for the actantial structure of a lexical unit;
- definition;
- paradigmatic and syntagmatic links between lexical units encoded with lexical functions: `Gener`(*carbon dioxide*) = *gas; greenhouse gas*;
- lexicographic examples to illustrate each sense.

In the following sections, we present the definitions only, while the remaining lexicographic information can be accessed through a visualization tool called *Spiderlex*,[17] which facilitates exploration of the data stored within the English Lexical Network (Ollinger *et al.* 2020).

[17] https://lexical-systems.atilf.fr/spiderlex/

Figure 5.8 illustrates the main elements of a definition formulated in accordance with the principles of the Explanatory and Combinatorial Lexicology. A definition consists of a ***propositional form***, which represents the actantial structure of a predicative lexical unit, along with two main elements: the obligatory ***central component*** and optional ***peripheral component(s)*** (Mel'čuk and Polguère 2018). The central component represents a minimal and generic paraphrase of the lexical unit being defined, while the peripheral components provide additional semantic characteristics that distinguish the given lexical unit from related ones.

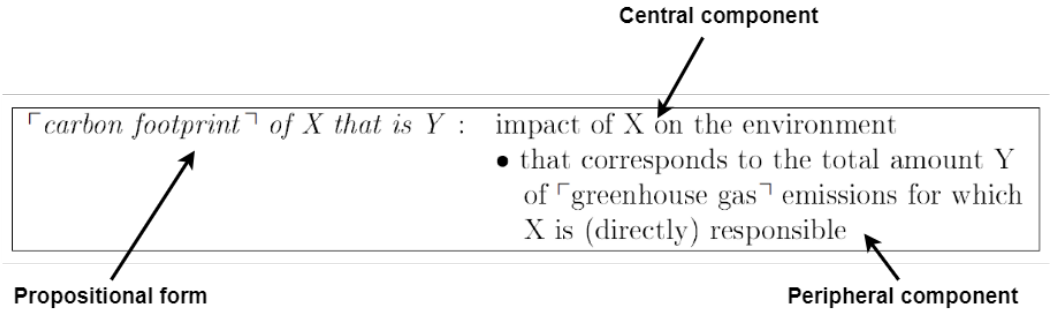

Figure 5.8 – The main elements of a definition that conforms to the principles of the Explanatory and Combinatorial Lexicology.

5.5.2 CARBON

5.5.2.1 CARBON **I.1**

CARBON **I.1** is the basic lexical unit of the vocable CARBON and a chemical Term used to denote the sixth chemical element in the Periodic table.[18] From a linguistic perspective, this terminological sense is not as straightforward as it may appear. In order to properly define CARBON **I.1**, it is necessary to refer to the notion of ***element***. Mikhel (2022) investigated the historical development of the notion of *element* and concluded that there is an ongoing debate among chemists on the proper way of defining *element* due to different perceptions of this notion. Specifically, there are two common perceptions which correspond to two senses of the vocable ELEMENT developed by Mikhel (2022) in the English Lexical Network:

[18]The Periodic table organizes chemical elements and provides information about their atomic mass, chemical symbol, and more.

- **spec** ELEMENT **III.3a** 'type of atoms **I.2**' [*He then rearranged the elements in the periodic table on the basis of atomic numbers.*]
- **spec critic** ELEMENT **III.3b** '⌜pure substance⌝' [*Approximately three-quarters of all known chemical elements are metals.*][19]

ELEMENT **III.3a** denotes an abstract entity that exists as a nominal element in the nomenclature of chemical elements, i.e., the Periodic table. In contrast, ELEMENT **III.3b** denotes a physical entity that is composed of atoms of ELEMENT **III.3a** and possesses physical characteristics, i.e., a pure substance. For example, oxygen (O) is an element with atomic number 8. The pure substance oxygen with the chemical formula O_2 possesses physical characteristics of a colorless gas. In chemistry, the preferred designation for oxygen (O_2) is *pure substance* and not *element*. Hence, the usage note **spec critic** (criticizable) for ELEMENT **III.3b** is used. Nevertheless, some lexicographic sources provide definitions of elements which combine abstract and physical characteristics. In (11) and (12), the element sulphur is defined with characteristics, such as *yellow*, *combustible*, *nonmetallic*, *pale*, that bring us to the physical world.

(11) The chemical element of atomic number 16, a **yellow combustible nonmetal**. [*Oxford Dictionaries*, https://www.lexico.com/definition/sulphur, 22/04/2020]

(12) A **pale yellow nonmetallic** element occuring widely in nature in several free, allotropic and crystal forms and combined in numerous sulfates and sulfides. [*The American Heritage Dictionary*, https://www.ahdictionary.com/word/search.html?q=sulfur, 22/04/2020.]

Taking this terminological confusion into consideration, we propose the following definition of CARBON **I.1** 'element' (see Figure 5.9). After extensive discussions on what information should be included in the definition of CARBON **I.1**, we have concluded that only essential characteristics of a given chemical element should be considered. It was a complicated task, considering that each element can be potentially defined with a wide range of properties, including wrongly defined with physical properties, e.g., abundance. Therefore, we thoroughly considered which

[19]Example sources: 1) https://www.asbmb.org/asbmb-today/science/020721/a-brief-history-of-the-periodic-table, 2) https://www.britannica.com/science/metal-chemistry.

properties should be included in the definition, and what should be left out as nonessential information. We considered it important to maintain a reasonable balance between a concise and informative definition, as opposed to encyclopedic definition.[20] For this, we identified the most basic and essential characteristic of a chemical element: its *atomic number*. Each element is uniquely associated with a specific atomic number which differentiates it from other elements. The atomic number of an element corresponds to the number of protons in the atomic nucleus (Law and Rennie 2020) and determines the element's position in the Periodic table.

carbon **I.1** :	element **III.3a** • with ⌜atomic number⌝ 6

Figure 5.9 – Definition of **spec** CARBON **I.1**.

Furthermore, CARBON **I.1** is labeled with the usage note **spec**, which indicates that it represents a full Term (see 3.3) predominantly used by experts in chemistry, physics, etc. While the name *carbon*, denoting a chemical element, can be familiar to non-experts, it is rarely used in daily discussions.

Examples.

(13) **Carbon** is the chemical element in the periodic table that has the symbol C and atomic number 6.
[Web, https://www.worldofmolecules.com/elements/carbon.htm, 15/07/2023]

(14) Graphene is a sheet of **carbon** atoms bonded to each other, just one atom thick; graphite is just an agglomeration of these sheets layered on top of each other.
[COCA, *Scientists identify hundreds of atomically-thin materials*, Ars Technica, 2018]

5.5.2.2 CARBON **I.2**

It is a common practice in chemistry to use the same designation for an element and a pure substance composed of atoms of this element. To reuse the previously mentioned example, *oxygen* is used to designate both the name of the element (O) and the name of the pure substance (O_2). However, the designation *carbon* is a notable exception to this rule: unlike other element names, *carbon* can only be

[20] Encyclopedic definition includes non-linguistic knowledge and associations with a given lexical unit. For instance, *was discovered in prehistoric times* is a piece of encyclopedic information about CARBON **I.1**.

used to refer to a chemical element and not a pure substance composed of carbon atoms. According to the terminological recommendation of the *IUPAC* (Fitzer *et al.* 1995), pure, or near pure, subtances composed of carbon atoms should be designated with specific Terms, such as *activated carbon*, *diamond*, *coke*, *carbon fibre*, *carbon black*, etc. In other words, the name *carbon* is only reserved to refer to the chemical element C.

Despite this terminological convention, the name *carbon* is used to refer to substances in both general and specialized texts. Presumably, experts use it because it resonates with their scientific practice, even though it might not comply with the recommended terminology. Furthermore, these subtle terminological nuances are not accessible to the general public with no background in chemistry.

We propose the following definition of CARBON **I.2** (see Figure 5.10). We consider CARBON **I.2** a runaway Term which is criticized in chemical terminology, hence the usage note **(spec critic)**. As non-experts are not aware of the duality of the notion of *element* (see 5.5.2.1), they tend to primarily associate the name *carbon* with a black solid substance. Hence, we consider CARBON **I.1** 'chemical element' a full Term and CARBON **I.2** a runaway Term.

carbon **I.2** :	substance **I.1a** • that is the materialization of carbon **I.1**

Figure 5.10 – Definition of **(spec critic)** CARBON **I.2**.

Examples.

(15) **Carbon** is a solid, with a color of blackish brownish resembling charcoal.
[Web, https://sites.google.com/a/mcsdonline.org/element-central/non-metals/carbon, 13/06/2023]

(16) In its diamond allotrope, **carbon** is the hardest known substance in nature.
[Web, https://www.ducksters.com/science/chemistry/carbon.php, 13/06/2023]

5.5.2.3 CARBON **II.1**

In sections 5.4, we outlined the main linguistic factors that shape the perception of *carbon* in the environmental discourse. We concluded that when used in the context of greenhouse gas emissions, *carbon* is an environmental Term which refers

to gases containing carbon atoms (CARBON **I.1**) that contribute to the greenhouse effect, e.g., carbon dioxide and methane. Additionally, *carbon* can also refer to carbon monoxide which is not considered a greenhouse gas. This perception of *carbon* corresponds to the sense CARBON **II.1** in the English Lexical Network (see Figure 5.11).

carbon **II.1** *released by X* :	gas **I.2** • that contains carbon **I.1** atoms **I.2** • that is released into the Earth's atmosphere by X • that is harmful for the environment when released in excess

Figure 5.11 – Definition of **(spec critic)** CARBON **II.1**.

In this definition, *X* is the actant slot which corresponds to the emitter of CARBON **II.1**, such as an activity (*volcanic activity*), a country, a factory, etc. We deliberately define CARBON **II.1** with the central semantic component *gas* and not *greenhouse gas* to account for cases when CARBON **II.1** denotes carbon monoxide which is not a greenhouse gas. Furthermore, the definition allows for interpretations of CARBON **II.1** as an individual gas (carbon dioxide), or a number of gases (carbon dioxide and methane).

The usage note **(spec critic)** signifies that it is a runaway Term criticized within specialized discourse. Researchers and educators emphasize the necessity of using precise names for the specific greenhouse gases involved, rather than the ambiguous umbrella expression *carbon*.

CARBON **II.1** has been of great importance in environmental talks and debates during recent decades. It is omnipresent in official reports of international environmental oragnizations as well as in scientific texts on environmental issues. It is a conspicuous example of a Term that has not been standardized or officially acknowledged yet and, thus, causes terminological confusion. It is reflected in the way the definition of CARBON **II.1** varies from source to source, with some sources neglecting it at all (see 5.4.3).

Examples.

(17) Coal-fired power plants, which produce the majority of Georgia's electricity and emit the most **carbon**, would pay the most. [COCA, Shelton Stacy, Halicks Richard,

Election 2008: The Challenge Of Change, 9 november 2008, Atlanta Journal Constitution]

(18) Across the globe, there are as many as 1.5 billion people without access to power, spending 40 to 70 per cent of their income on kerosene and firewood, with two million deaths a year from smoke inhalation and 150 million tonnes of **carbon** released annually.
[COCA, Johnson Leo, Petropolis now, 11 august 2013, New Statesman, Vol. 142 Issue 5183, p. 24-28]

5.5.2.4 CARBON **II.2**

In section 6.3, we will discuss the frames that determine the way *carbon* is shaped in environmental discourse, i.e., ideological coloring and abstraction. Abstraction consists in perceiving carbon as an intangible and invisible commodity that is being traded in carbon markets. Additionally, abstraction results from the type of motivational environmental discourse which encourages individuals to assess, control and reduce **their** emissions. As a result, carbon has become a nonmaterial by-product of one's unsustainable lifestyle. Linguistically, this is expressed by constructions which involve personal pronouns and the use of peculiar verbs, e.g., *my carbon*, *she uses too much carbon*, *they emit carbon*, etc.

What is exactly meant by these expressions? Clearly, people do not physically emit carbon in the literal sense.[21] Instead, they indirectly contribute to carbon emissions by buying products that were produced at emitting factories and transported using emitting transport, or by taking flights that generate emissions. In other words, these expressions refer to an individual's lifestyle and its contribution to carbon emissions – a sort of logical shorcut.

In the definition (see Figure 5.12), the component 'symbolic substance' points to its figurative nature, as it is a symbol of pollution but not the polluting substance itself. We consider CARBON **II.2** an emblematic example of a quasi-Term (labeled with the usage note **quasi-spec**), as lay people use it in every day speech while referring to something belonging to environmental science as if it were a Term while not being scientific in their reasoning.

[21] In fact, people do emit "carbon" in form of carbon dioxide during respiration process. However, this is not the reference in the mentioned expressions.

carbon **II.2** *produced by X* :	symbolic substance • that represents the amount of polluting substance released by X [as if it were carbon **II.1**]

Figure 5.12 – Definition of **quasi-spec** CARBON **II.2**.

In terms of how CARBON **II.2** is lexicalized in the language, it is a metaphor derived from CARBON **II.1**. However, lay Speakers do not necessarily realize the metaphoric aspect and often use CARBON **II.2** as if it were the Term CARBON **II.1**. In 5.12, the lexicographic comment found between square brackets – *as if it were carbon **II.1*** – does not serve to add semantic value to the definition but rather indicates the position of this sense within the polysemy of the vocable CARBON.

Examples.

(19) Most people emit **carbon** every day simply by using a non-renewable resource, such as coal, natural gas, or oil.
[Web, https://www.wise-geek.com/what-is-a-carbon-footprint-test.htm, 13/06/2023]

(20) You can find green wedding DJs or green photographers or videographers, who already reduce their **carbon** use by using low-powered high-output equipment and offset any remaining carbon use.
[Web, http://www.wikihow.com/Have-a-Green-Wedding, 13/06/2023]

5.5.3 ⌜CARBON DIOXIDE⌝

The terminological vocable ⌜CARBON DIOXIDE⌝ comprises three lexical units:

- **spec** ⌜CARBON DIOXIDE⌝ **1** 'type of molecule **I**'
- **(spec)** ⌜CARBON DIOXIDE⌝ **2** 'gas **I.2** that is a component of the Earth's atmosphere'
- **spec** ⌜CARBON DIOXIDE⌝ **3** 'gas **I.2** that is used for industrial purposes'

5.5.3.1 ⌜CARBON DIOXIDE⌝ **1**

⌜CARBON DIOXIDE⌝ **1** is a chemical Term which refers to a type of molecules with the chemical formula CO_2 (see Figure 5.13). Molecules are chemical entities composed of atoms and represent the smallest units of substances (Mikhel 2022: 184). For example, a carbon dioxide molecule consists of one carbon atom and two

oxygen atoms, and it serves as the smallest unit of carbon dioxide gas. While the atomic number is the most important defining characteristic for chemical elements, for molecules, it is their atomic composition reflected in their chemical formula.

⌜*carbon dioxide*⌝ **1** :	type of molecule **I** • that has the formula CO_2

Figure 5.13 – Definition of **spec** ⌜CARBON DIOXIDE⌝ **1**.

When it came to defining ⌜CARBON DIOXIDE⌝ **1**, we encountered two options. Firstly, we could define a molecule by enumerating all the relevant atoms, or secondly, we could provide the chemical formula. We opted for the latter approach since some molecules can be quite complex – for example, barium acetate molecule has the formula $Ba(C_2H_3O_2)_2$, making it impractical to define them by listing all the atoms contained within the giving molecule. Hence, we provide the chemical formula of carbon dioxide, i.e., CO_2, in the definition.

Examples.

(21) The level of **carbon dioxide** determines the pH of the blood.
[COCA, WALMSLEY Sandy, *Oxygen and blood gas sampling in the community*, 2 november 2011, Practice Nurse, Vol.41 Issue 2, p.41-45]

(22) **Carbon dioxide** has the formula CO_2 and at the centre of this linear molecule is a carbon atom joined by two pairs of double-bonds to the oxygen atoms, i.e O=C=O.
[Web, http://www.chm.bris.ac.uk/motm/CO2/CO2jm.htm, 07/06/2023]

5.5.3.2 ⌜CARBON DIOXIDE⌝ **2**

CARBON DIOXIDE **2** denotes a gas that is a natural component of the Earth's atmosphere. In the definition, we use the strong semantic component *that is a component of the Earth's atmosphere* and the weak semantic component[22] *which is harmful for the environment when released in excess* to refer to the dual perception of carbon dioxide in environmental discourse (see Figure 5.14). By formulating the definition this way, we model the fact that CARBON DIOXIDE **2** is first and foremost a natural constituent of the Earth's atmosphere and its quality of being harmful

[22]Weak components are optional components in definitions, as opposed to "strong", or obligatory, components (Mel'čuk and Polguère 2018: 13).

for the environment is optional as it becomes relevant when the concentration of CARBON DIOXIDE **2** is too high.

⌜*carbon dioxide*⌝ **2** *released by X* :	⌜carbon dioxide⌝ **1** in form of a gas **I.2** • that is a component of the Earth's atmosphere • that is released into it by X • that is harmful for the environment when released in excess

Figure 5.14 – Definition of **(spec)** ⌜CARBON DIOXIDE⌝ **2**.

The actant *X* implies various activities on Earth that involve both natural and anthropogenic causes. These activities can include volcanic eruptions, respiratory processes, industrial emissions, and more.

The usage note **(spec)** indicates that it is a runaway Term which has migrated from specialized environmental discourse on carbon dioxide emissions and has firmly taken its place in the public debate on environment issues.

The lexical unit CARBON DIOXIDE **2** is the specialization of the basic lexical unit CARBON DIOXIDE **1** (encoded with copolysemy relation *Specialization*). It implies that CARBON DIOXIDE **2** is a more specific sense with additional semantic nuances, while CARBON DIOXIDE **1** serves as a generic lexical unit and is used in the central component of the definition of CARBON DIOXIDE **2**.

Examples.

(23) The combustion of fossil fuels drives the world's energy production, but it also emits **carbon dioxide** (**CO_2**) and other greenhouse gases.
[COCA, *Crucial role of electrical engineers in addressing climate change*, 24 march 2016, Phys.Org]

(24) Currently, the best proposed way to remove **carbon dioxide** is through direct air capture, which involves pumping air through a system that liquefies and stores the **carbon dioxide** or converts it into a substance that is either inert or useful.
[Web, https://www.arcticwwf.org/the-circle/stories/can-carbon-removal-save-us/, 07/06/2023]

5.5.3.3 ⌜CARBON DIOXIDE⌝ 3

Unlike CARBON DIOXIDE **2** that refers to "unintentional" by-product of natural and industrial processes, CARBON DIOXIDE **3** denotes a synthesized and commer-

cialized product manufactured and used for specific purposes. These purposes include beverage carbonation, fire suppression, food preservation, decaffeination (25), welding (26), and more.

In the definition (Figure 5.15), the actans X and Y stand for the agent and the application of carbon dioxide gas, respectively. For example, in *the company uses carbon dioxide to create carbonated beverages*, X is realized by *the company* and Y is realized by *create carbonated beverages*.

⌜*carbon dioxide*⌝ **3** *used by X for Y* :	⌜carbon dioxide⌝ **1** in form of a gas **I.2** • that is industrially produced to be used by X for Y

Figure 5.15 – Definition of **spec** ⌜CARBON DIOXIDE⌝ **3**.

The technical nature of the sense CARBON DIOXIDE **3** makes it a full Term (labeled with the usage note **spec**) that is not commonly used in every day lay conversations.

Similarly to CARBON DIOXIDE **2**, the lexical unit CARBON DIOXIDE **3** is the specialization sense of the basic lexical unit – CARBON DIOXIDE **1** (encoded with copolysemy relation *Specialization*). The polysemy structure of the vocable CARBON DIOXIDE is illustrated in Figure 5.16.

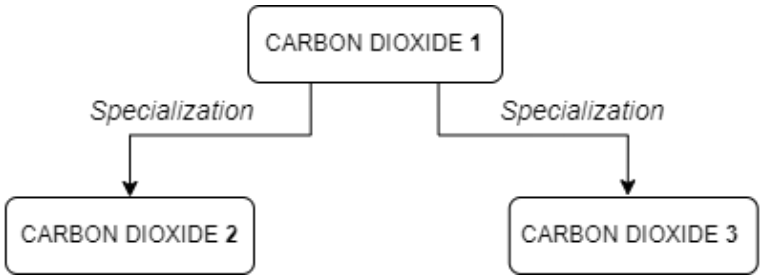

Figure 5.16 – Polysemy structure of the vocable CARBON DIOXIDE.

Examples.

(25) During the decaffeination process, the **carbon dioxide** moves through and extractor and a scrubber, on a loop, for approximately 8-9 hours.
[Web, https://www.futurelearn.com/info/courses/everyday-chemistry/0/steps/22338, 07/06/2023]

(26) **Carbon dioxide**, in combination with other gases, also helps to enhance the toughness of weldment.
[Web, https://www.austgen.com.au/why-is-co2-used-in-welding/, 07/06/2023]

5.5.4 ⌜CARBON CYCLE⌝

The terminological vocable ⌜CARBON CYCLE⌝ comprises two lexical units:

- **spec** ⌜CARBON CYCLE⌝ **1** 'cyclic displacement of carbon **I.1** atoms **I.2**'
- **spec** ⌜CARBON CYCLE⌝ **2** 'cycle of thermonuclear reactions in stars'

Unlike *carbon footprint*, which will be discussed in the next section, *carbon cycle* is rarely included in educational "green" glossaries aimed at explaining "carbon jargon." Knowledge of the carbon cycle is crucial in understanding how Earth's climate changes and in developing effective carbon management programs (Dilling *et al.* 2003; NASA 2015). However, carbon cycle is a "subject of confusion" among the general public (Mackey *et al.* 2013), as often educational programs on carbon do not address the basics of chemistry and physics relevant to the global carbon cycle (McCaffrey and S. M. Buhr 2008).

Carbon cycle describes a natural process which is characterized by the continuous movement of carbon atoms contained in substances such as carbon dioxide (CO_2), carbonic acid (H_2CO_3), glucose ($C_6H_{12}O_6$), calcium carbonate ($CaCO_3$), and more. These substance undergo transformation by converting into one another and ensure the movement of carbon atoms.

While carbon cycle is primarily associated with the planet Earth, research has shown that this process is also relevant to other celestial bodies. However, these discoveries are still in progress, and experts are in the early stages of defining the carbon cycle of Mars.[23] The carbon cycle on the Earth, on the other hand, has been relatively well studied. It is acknowledged that the displacement and transformation of carbon-containing substances between systems of the Earth (such as the atmosphere, biosphere, lithosphere, and so on) is one of the most crucial natural processes on the planet.

We made an observation that in educational sources, carbon cycle is often depicted with a specific emphasis on carbon dioxide. While carbon dioxide plays a significant role in the carbon cycle, it is not the sole participant, and other

[23] https://www.nasa.gov/feature/goddard/2022/nasa-s-curiosity-rover-measures-intriguing-carbon-signature-on-mars

carbon-containing substances should be considered as well to avoid promoting a reductionist perception of the carbon cycle. Hence, we propose the definition (see Figure 5.17) which describes ⌜CARBON CYCLE⌝ **1** as a 'cyclic displacement and transformation of substances containing carbon atoms'. The actant *X* stands for a celestial body that has a carbon cycle.

⌜*carbon cycle*⌝ **1** *of X* :	biogeochemical process • that consists in cyclic displacement and transformation of substances **I.1b** containing carbon **I.1** atoms **I.2** between systems of a celestial body X

Figure 5.17 – Definition of **spec** ⌜CARBON CYCLE⌝ **1**.

Additionally, we considered including the semantic component *that can be disrupted by human activity* in the definition, but we eventually excluded this option. While it is true that in the environmental context, excessive greenhouse gas emissions are associated with the potential carbon cycle disbalance, the concept of the carbon cycle itself does entail an anthropocentric aspect that would be necessary for a comprehensive definition.

Although the Term *carbon cycle* does appear in mass media discussions on the disruption of the carbon cycle caused by anthropogenic activities, we do not consider ⌜CARBON CYCLE⌝ **1** a runaway Term. This is because it is not an environmental buzzword of public talks on environmental issues. Furthermore, to our knowledge, when non-experts do use *carbon cycle*, they refer mainly to carbon dioxide, which resonates with the one-sided interpretation of the concept provided in educational sources and the media. Further research is necessary to determine whether such distorted public vision of the carbon cycle is really common.

Examples.

(27) A recent study of the global **carbon cycle** offers a new perspective of Earth's climate records through time.
[COCA, *Fossil algae reveal 500 million years of climate change*, 28 november 2018, Phys.Org]

(28) Nitrogen is now understood to help regulate the **carbon cycle** and exert both cooling and warming effects on the climate.
[COCA, Lougheed Tim, *Reactive nitrogen wrangling strategies for mitigating pollution*, 2012, Environmental Health Perspectives, Vol. 120 Issue 5]

The sense ⌜CARBON CYCLE⌝ **2** is an astronomy Term that refers to the cycle of reactions occuring in stars. As ⌜CARBON CYCLE⌝ **2** is semantically distant from ⌜CARBON CYCLE⌝ **1** and is not pertinent to the environmental discourse, we did not provide any rigorous lexicographic analysis for it in the English Lexical Network.

5.5.5 ⌜CARBON FOOTPRINT⌝

The notion of *carbon footprint* was first developed in the 1990s as a symbolic representation of the human impact on the Earth (Niranjan 2022). In 2004-2006, the British oil and gas company BP launched a media campaign[24] that popularized the Term *carbon footprint*. Many believe that this campaign aimed to shift environmental responsibility onto consumers and their individual actions (Supran and Oreskes 2021). In recognition of its significance, the *Oxford English Dictionary* named *carbon footprint* as a word of the year in 2007 (Oxford Languages 2023).

The calculation of carbon footprint can concern a person, an entity (such as a country, a company, etc.), or a process to estimate how much greenhouse gases have been emitted taking into consideration different criteria. For example, personal carbon footprint would include criteria such as lifestyle choices concerning energy, diet, travel and other consumer choices.

Carbon footprint is quantified in mass units called carbon dioxide equivalent (CO_2-eq). This metric was developed to enable the comparison of emissions from different greenhouse gases based on their global warming potential, using the greenhouse gas carbon dioxide as a reference (EEA 2001). Despite efforts to standardize the measurement and reporting of carbon footprint (ISO 2018), the concept remains vague and complex, with the lack of clear standards, methodologies, regulations, and the challenge of very complex calculations (Markey 2023). As a result, carbon footprint calculations often provide an incomplete picture and fail to account for the full extent of emissions (Berners-Lee and Clark 2010). Additionally, due to the complexity of the concept, carbon fooptrint is often misunderstood by non-specialists. For example, 59% (n = 1000) of businessmen were confused

[24]The advertisement video is accessible on YouTube:
https://www.youtube.com/watch?v=ywrZPypqSB4&ab_channel=mpdrsn.

about its meaning (Moore 2022).

Questions have arisen regarding the appropriatness of the designation *carbon footprint*. Alternative designations such as *carbon weight* (Hammond 2007) and *climate footprint* Wiedmann and Minx (2007) have been suggested to emphasize the measurement in mass units (kilograms or tonnes) or to refer to the most comprehensive account of all the greenhouse gases.

While carbon footprint measurements commonly take into account multiple greenhouse gases, the Intergovernmental Panel on Climate Change (IPCC 2023b) borrows the definition from Wiedmann and Minx (2007) who defined *carbon footprint* with a focus on carbon dioxide emissions:

> Measure of the exclusive total amount of emissions of carbon dioxide (CO_2) that is directly and indirectly caused by an activity or accumulated over the life stages of a product.

In reality, calculations of one's carbon footprint may include greenhouse gases other than carbon dioxide, e.g., methane and nitrous oxide. For example, Yao *et al.* (2022) studied the contribution of nitrous oxide emissions to the carbon footprint of wastewater treatment plants.

In our definition (Figure 5.18), we point out that ⌜CARBON FOOTPRINT⌝ has two actant positions: *X* stands for an agent responsible for emissions and *Y* stands for the value in units of carbon dioxide equivalent, which corresponds to the amount of emissions. Furthermore, we draw a semantic link with ⌜GREENHOUSE GAS⌝ and not with CARBON **II.1** because the calculations of carbon footprint can take into account non-carbon-containing greenhouse gases, e.g., nitrous oxide. Finally, we introduce a weak semantic component 'directly' to indicate that *X* can be directly or indirectly responsible for the given emissions.

⌜*carbon footprint*⌝ *of X that is Y* :	impact of X on the environment • that corresponds to the total amount Y of ⌜greenhouse gas⌝ emissions for which X is (directly) responsible

Figure 5.18 – Definition of **(spec)** ⌜CARBON FOOTPRINT ⌝.

We consider ⌜CARBON FOOTPRINT⌝ a runaway Term, and hence, we label it with the usage note **(spec)**. Unlike ⌜CARBON CYCLE⌝ **1** (that we labeled as **spec**), the Term ⌜CARBON FOOTPRINT⌝ does feature in public talks on environmental issues. This is due to the fact that the concept of carbon footprint itself is closely linked to public efforts in environmental mitigation.

Examples.

(29) When burned in a power plant, natural gas has a smaller **carbon footprint** than coal, and when it displaces coal, emissions decline.
[COCA, WALD L. Matthew, *The Potential Downside of Natural Gas*, 2014, New York Times]

(30) The choices we make in our homes, our travel, the food we eat, and what we buy and throw away all influence our **carbon footprint** and can help ensure a stable climate for future generations.
[COCA, *Free Carbon Footprint Calculator*, 2012, http://www.nature.org/greenliving/carboncalculator/index.htm]

5.5.6 ⌜CARBON CAPTURE AND STORAGE⌝

Nowadays, numerous technologies exist to address emissions reduction and management, including carbon capture and storage (CCS), carbon capture and use (CCU), carbon capture and use or sequestration (CCUS), direct air capture with carbon storage (DACCS), carbon dioxide removal (CDR), bioenergy with carbon capture and storage (BECCS), etc. The development of these technologies is actively underway, as well as the efforts to categorize, conceptualize and establish standardized terminology. Educators point out that there is an unnecessary complex interplay of notions related to carbon management technologies which can hamper the dialog concerning the development, public acceptance, and utilization of such technologies (Olfe-Kräutlein *et al.* 2022). For example, the acronym CDR stands for both *carbon dioxide recycling* and *carbon dioxide removal* which can cause potential misunderstandings when used in a single context.

Leaving aside the general terminological confusion concerning low-carbon technologies, we focus on *carbon capture and storage* technology, which refers to the process of capturing carbon dioxide from industrial facilities, transporting it via pipelines to storage locations and isolating it permanently deep underground or in

the ocean (Metz *et al.* 2005). With almost 200 CCS projects currently in progress in the world, this technology is a promising asset in the field of environmental mitigation. The European Commission recognizes carbon capture and storage technology as "an important part of the EU decarbonisation effort" (European Commission 2023a).

There are two accepted variants of the names of this technology: *carbon dioxide capture and storage* and the shorthand name *carbon capture and storage*. According to terminological recommendations, it is preferable to use the full form *carbon dioxide capture and storage*, as it explicitly points to carbon dioxide, the gas that is the object of capture and storage. However, the shorthand form *carbon capture and storage* is commonly used in scientific literature and is acknowledged by official environmental organization, such as the *International Energy Agency*, the *Intergovernmental Panel on Climate Change*, and the *United Nations Framework Convention on Climate Change*, and inscribed in organisms' official names, e.g., the *UK Carbon Capture and Storage Research Community*. As with *carbon dioxide emissions* and *carbon emissions* (see 5.4.1), the shorthand version *carbon capture and storage* seems to prevail. We performed a search in Ngram viewer which revealed that the shorthand form *carbon capture and storage* is prevalent over *carbon dioxide capture and storage* (Figure 5.19). Furthermore, we performed the same search in our specialized corpus V2 (for the discussion on corpus, see 4.1.3) which showed that the form *carbon capture and storage* form has 20 occurrences, while *carbon dioxide capture and storage* has 6 occurrences.

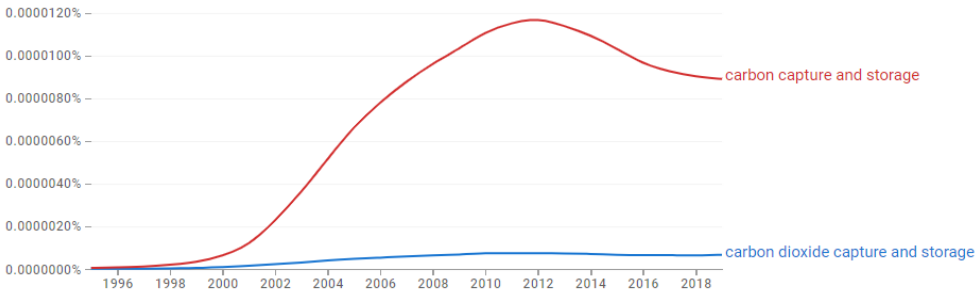

Figure 5.19 – Ngram Viewer search for *carbon dioxide capture and storage* and *carbon capture and storage* in the English corpus.

The formal definition (see Figure 5.20) is straightforward and points out essential actions performed with carbon dioxide in the process of carbon capture

and storage. ⌜CARBON CAPTURE AND STORAGE⌝ is a full Term (usage note **spec**) which is predominantly used by environmental and industrial experts.

⌜*carbon capture and storage*⌝ *by X* :	industrial process performed by X • that consists in collecting and storing emissions of ⌜carbon dioxide⌝ **2** before it is released into the atmosphere

Figure 5.20 – Definition of **spec** ⌜CARBON CAPTURE AND STORAGE⌝.

Examples.

(31) Negative net carbon emissions could in principle be achieved by widespread use of **carbon capture and storage** technology, driven by non-carbon energy sources.
[COCA, GREAVES Hilary, *Climate Change and Optimum Population*, 2019, The Monist]

(32) To solve this problem, companies say they will employ carbon sequestration or **carbon capture and storage** (CCS), in which the CO2 emissions are stored, usually deep underground, rather than released into the atmosphere.
[COCA, *Ridgeway* James, *Scrubbing king coal*, 2008, Vol. 33, Iss. 3, Mother Jones]

In conclusion, our lexicographic analysis of the three idioms, namely ⌜CARBON CYCLE⌝, ⌜CARBON FOOTPRINT⌝ and ⌜CARBON CAPTURE AND STORAGE⌝, revealed that *carbon* refers to different concepts in each of them (see Table 5.4). While we based our interpretations on commonly accepted scientific standards, it should be noted that the meaning of the idioms can vary with context. For example, in some contexts, *carbon* in ⌜*carbon cycle*⌝ refers to carbon dioxide, and in ⌜*carbon footprint*⌝, it is restricted to one greenhouse gas only. This variability in the meaning of *carbon* within idioms hinders the intuitive intepretation of the *carbon* lexicon. As a result, it leads to confusion among the general public and obstructs effective communication between experts.

Idiom	Interpretation of *carbon*
⌜CARBON CYCLE⌝ **1**	'chemical element C'
⌜CARBON FOOTPRINT⌝	'greenhouse gases'
⌜CARBON CAPTURE AND STORAGE⌝	'carbon dioxide'

Table 5.4 – *Carbon* idioms and their respective interpretations.

5.6 Chapter synthesis

In this chapter, we discussed the notion of *carbon* from multiple perspectives. First, we explored its role in chemistry and environmental science. The recognition of carbon as a chemical element dates back to 1789. At present, the name *carbon* is used in various contexts, and it has become ubiquitous in environmental discourse primarily due to its association with the dominant greenhouse gas, carbon dioxide. However, such popularity coexists with an apparent lack of clear and consistent terminological description of *carbon* in specialized glossaries. Due to this lack of clarity, the general public often experiences confusion as to what *carbon* denotes.

Furthermore, we analyzed the linguistic features of *carbon* as it is used in both specialized and press discourse, with a specific focus on the reasons contributing to terminological confusion. These reasons include the context-dependent variation in meaning, its marginal character in specialized discourse, and the resulting definitional lacuna in specialized glossaries.

We presented a lexicographic description of the vocables CARBON, ⌜CARBON DIOXIDE⌝, ⌜CARBON CYCLE⌝, ⌜CARBON FOOTPRINT⌝ and ⌜CARBON CAPTURE AND STORAGE⌝ in the English Lexical Network. The choice of these particular vocables was motivated by the need to provide clear and formalized lexicographic information regarding the definitions of the corresponding lexemes and the semantic links between them.

Chapter 6

Conceptualization of *carbon* in ordinary discourse

Summary

To explore the general public's perception of *carbon* within our social network corpora, we performed *qualitative content analysis* (Krippendorff 2004; Schreier 2013) of retrieved online discussions. Krippendorff (2004: 410) provides the following definition of such analysis:

> [Qualitative content analysis] involves a close reading of textual matter, reorganizing relevant parts of it into analytical categories, and creating interpretations, narratives of scholarly interest relating to the meanings and uses of the analyzed Text.

Our qualitative analysis consisted in the study of texts, such as tweets and Reddit posts, by close reading in order to identify thematic patterns related to how users perceive, conceptualize and lexicalize *carbon*. In this analysis, we concentrate solely on the immediate and isolated context of the texts, disregarding the author's background and pragmatic intentions since such information is unavailable to us (see 4.2). Furthermore, close reading implies manual examination of texts, as opposed to NLP-based techniques. The advantages of employing manual discourse analysis include the ability to capture the nuanced information that may not be accessible through automated text processing, greater flexibility in understanding the texts and their specific characteristics.

We had three specific research questions in mind to guide us through the analysis:

1. How effective is the utilization of social network content for the purposes of qualitative linguistic analysis?
2. What are the perceptions and conceptualizations of *carbon* within ordinary discourse found on social networks?
3. What semantic links can be identified between *carbon* and other related expressions (*carbon dioxide*, *methane*) when they coexist within a single context?

The preliminary findings of this analysis were presented in Gotkova and Chepurnykh (2022). Although we have since made significant progress, we still incorporate certain statements from the article.

6.1 Building a subcorpus

Our main Twitter and Reddit corpora (see 4.4) contain a significant amount of data. The keyword `carbon` appears 161,773 times in Twitter data and 19,928 times in Reddit data. To effectively explore the usage of *carbon* and study its semantic links with other expressions, we created a subcorpus by filtering the corpora using specific keywords to create "Carbon subcorpus" (further in the text, referred to as *subcorpus*). The filtering condition required a tweet or a Reddit post (a comment or a submission) to include the keyword `carbon` along with at least one of the following keywords:

- `carbon dioxide`,
- `methane`,
- `co2` (the chemical formula of carbon dioxide, CO_2),
- `ch4` (the chemical formula of methane, CH_4).

As our research progressed, we found that in specialized texts, *carbon* can also refer to other (greenhouse) gases, such as carbon monoxide, nitrous oxide, sulphur hexafluoride, and more. Hence, we decided to supplement our subcorpus by extracting texts that contained keywords referring to such gases:

- `carbon monoxide`/`co`,
- `nitrous oxide`/`n20`,
- `hydrofluorocarbon(s)`/`hfc(s)`,
- `perfluorocarbon(s)`/`pfc(s)`,
- `sulphur hexafluoride`/`sf6`.

Although we assumed that the supplementary keywords might be too specific to be used in ordinary discourse, we wanted to be methodical and test our hypothesis with the most complete list of keywords.

As a result, we extracted 3,204 tweets and 1,591 Reddit posts (1,552 comments and 39 submissions) that were suitable for the qualitative analysis. The disbalance in the number of texts was corrected and even inverted by the difference in average length of texts in the two social networks (see 4.4). We obtained a total of 121,341

tokens in tweets and 339,639 tokens in Reddit posts. The average length of a tweet was 38 tokens, while an average Reddit post consisted of 218 tokens. In both corpora, the most frequent keywords were `carbon dioxide`, `co2` and `methane` (see Figure 6.1). This can be attributed to the fact that in scientific and mass media environmental context, *carbon* is primarily associated with carbon dioxide and methane. Hence, the public is more familiar with these greenhouse gases. The remaining keywords had low frequencies, with the lowest being:

- `hydrofluorocarbon(s)` (1 occurrence in Twitter; 2 in Reddit),
- `HFC(s)` (5; 6),
- `perfluorocarbon(s)` (1; 0),
- `PFC(s)` (1; 0),
- `sulphur hexafluoride` (0; 0),
- `SF6` (1; 1).

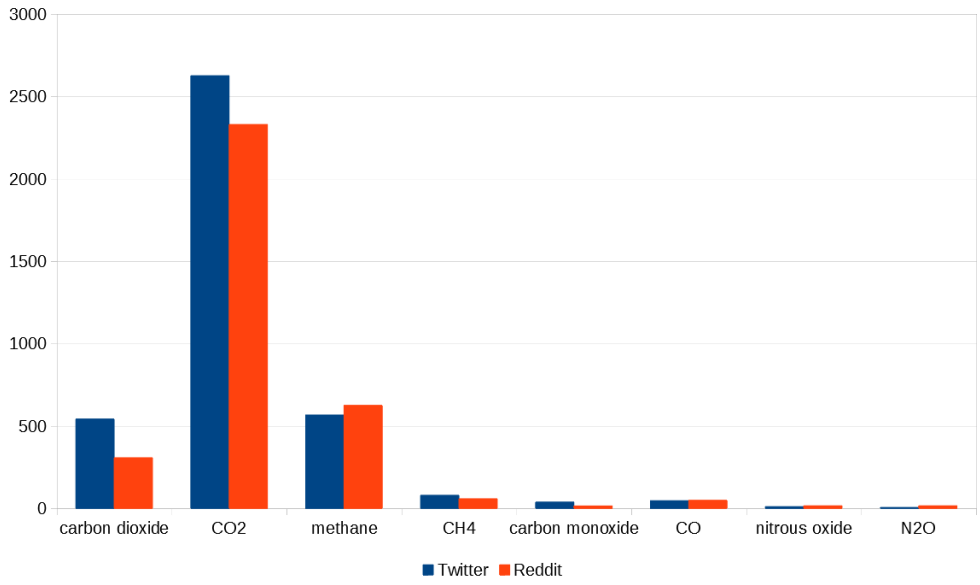

Figure 6.1 – The frequency of keywords in the corpus sample.

While reviewing the extracted texts, we identified some that were unsuitable for our qualitative analysis. For example, certain tweets lacked context for proper analysis due to their insufficient length. Furthermore, some texts displayed characteristics more pertinent to mass media and scientific discourse than to the ordinary discourse of the general public, as these were advertisements (1), and dissemination texts (2). Texts falling into these categories were marked as irrelevant and were not considered for the qualitative analysis.

(1) Learn about how human activity is impacting the carbon cycle and causing atmospheric carbon dioxide to increase. This #climatechange interactive lesson is ideal for students to understand global warming as they test variables on their own. [Twitter, ID: 1301112705177325570]

(2) A recent study shows that "countries with carbon prices on average have annual carbon dioxide emissions growth rates that are about two percentage points lower than countries without a carbon price, after taking many other factors into account." [Twitter, ID: 130119696707525017 8]

6.2 Conceptualizations of *carbon*

Close analysis of each text through multiple iterations allowed us to identify repeated thematic patterns that reflect how social networks users conceptualize *carbon* and how these conceptualizations are expressed in the texts (see Figure 6.2). As we processed the subcorpus, each text was categorized either as an illustration of a previously identified thematic pattern or served as the foundation for a new one. Consequently, each text received two distinctive labels. The first label indicates how *carbon* was conceptualized in the text, while the second label represents the semantic links with the keywords used to build the corpus sample (`carbon dioxide`, `co2`, `methane`, `ch4`, etc.).

C. I. *carbon* is a chemical element

C. II. *carbon* is a solid

C. III. *carbon* is a gas

C. III.1. *carbon* ≡ *carbon dioxide*

C. III.2. *carbon* ≢ *methane* (*carbon* ≡ *carbon dioxide*)

C. III.3. *carbon* ≡ *carbon dioxide*/*methane*

Figure 6.2 – **C**onceptualizations of *carbon* identified in the subcorpus.

We will now provide a description of each conceptualization, complemented with examples from our corpus. Although the analysis was based on a total of 4,795 texts, we will only present a limited number of examples either in full or abridged version. In examples, the standalone *carbon* or *carbon* idioms are given in bold and semantically related expressions are underlined.

While our primary focus was on the standalone use of *carbon*, our subcorpus contained numerous instances of *carbon* idioms, such as ⌜*carbon footprint*⌝, ⌜*carbon market*⌝, ⌜*carbon price*⌝, ⌜*carbon sequestration*⌝, etc. We cannot ignore such idioms due to their importance in environmental discourse. Hence, we considered them for the analysis as well. However, it should be noted that the meaning of *carbon* within an idiom does not necessarily correspond to the meaning of *carbon* as a standalone unit. This can be explained by the principle of *semantic non-compositionality* of idioms (Mel'čuk 2023: 30) – the meaning of a given idiom is not the sum of meanings of the idiom's components.

Conceptualization I: *carbon* is a chemical element. The first conceptualization of *carbon* in our subcorpus corresponds to the sense CARBON **I.1** we developed in the English Lexical Network (see 5.5.2.1), which refers to chemical element C. This conceptualization features chemical element carbon[1] as a constituent of chemical compounds such as carbon dioxide, methane, and carbon monoxide. The semantic relationship between *carbon* as a constituent and *methane* and *carbon dioxide* as chemical compounds containing carbon is characterized by a part-whole relationship, specifically a meronymic link.

Linguistic evidence that supports the conceptualization of *carbon* as a chemical element comes from the way the Term is used in the context of its chemical properties, the role in compounds, and interactions with other elements. In (3), the author emphasizes the difference in molecular weights between methane and carbon dioxide, specifically they say that it would require four times the amount of carbon to generate the same mass of methane as compared to generating the same mass of carbon dioxide. Example (4) illustrates the presence of carbon atoms in methane belched by cattle. In (5) and (6), the users mention how carbon interacts with oxygen in a living tree, and with oxygen and hydrogen to create water vapor and carbon dioxide.

(3) Methane is 30x more potent as a greenhouse gas than CO2 **by mass**. i.e. 1 kg of methane has the same global warming impact as 30kg of CO2, but methane is

[1] In correct chemistry terms, we should rather say *carbon atoms* (as discussed in 5.5.2.1). However, in our subcorpus, the users do not differentiate between chemical element as a nominal unit of the Periodic table and atoms as physical constituents of molecules and compounds.

significantly lighter (about 4x) than CO2, so it takes roughly 4x as much **carbon** to produce the same mass of methane compared to CO2. [Reddit, ID: e7nq7rs]

(4) Methane from beef is offset by the CO2 consumed by the plants/grass they eat. This is just part of the natural atmospheric carbon cycle. Those **carbon atoms** that cattle belch, in the form of methane, need to come from somewhere (the atmosphere). [Reddit, ID: eoe8rqx]

(5) Planting new trees every year isn't a bad thing though. Younger trees and saplings are really good at capturing CO2 because a tree is made of **carbon** and oxygen, and growing trees hold into the CO2 very well. [Reddit, ID: gf002ed]

(6) So like methane combusts right? And then the hydrogen and **carbon** synthesizes with the oxygen in the air creating water vapor (H2O) and Carbon Dixoide (CO2), you know? [Twitter, ID: 1301619664460505088]

This conceptualization is relevant to the chemistry aspect of environmental discussions. While our analysis of the subcorpus has not revealed any significant deviations from the specialized usage (CARBON **I.1**), we acknowledge that the ambiguity created by the use of *carbon* 'chemical element' and *carbon* 'carbon dioxide' in one context leads to confusion among the general public. This will be further discussed in the paragraph on Conceptualization III.1 and Section 6.3.4.

Conceptualization II: *carbon* is a solid. The way *carbon* is conceptualized as a solid substance in the subcorpus loosely corresponds to the sense CARBON **I.2** (see 5.5.2.2) in the English Lexical Network, which refers to a pure substance made of carbon atoms. Although criticized in chemistry terminology, it is a common practice, even among specialists, to refer to pure carbon-containing substances as *carbon*. However, in the ordinary discourse, the solid carbon is rather associated with coal or other black substances which are not necessarily pure forms of carbon.

The distinctive linguistic elements that allowed us to identify texts that contribute to this particular conceptualization of *carbon* included the physical characteristics attributed to *carbon* and its contextual associations with other related Terms such as *graphite* and *carbon dioxide*. For example, in (7), the use of *solid graphite* implies a solid form of carbon. In (8), *carbon* is referred to as *physical dirt* and *chunks* that can be buried as waste products of carbon dioxide removal. Similarly, in (9), carbon is viewed as a *black* substance meant to be mixed with

soil. Finally, in (10), the author contrasts *solid carbon* and *gaseous carbon dioxide* to emphasize the difference in their properties.

(7) If you could make solid graphite, basically pure **carbon**, you could probably put it in some open pit mines and get rid of it. [Reddit, ID: f5lp08y]

(8) My honest solution is to (somehow) turn the CO2 into physical **carbon** dirt, chunks, whatever, and bury it. And keep buying it. Or blast it into space. I presume you can't drop it into the ocean because it would do what we are already doing to the oceans – acidification. So my proposal: dump it in a desert, cover it with sand. [Reddit, ID: g9dgyvi]

(9) Soil should be black with **carbon** and be feet or meters deep. Instead they are anemic because the **carbon** has all been blow off, into the atmosphere as CO2 or washed away into the nearest river. [Reddit, ID: enmx7tn]

(10) I'm sure you'll recall chemistry at school that solid **carbon** and gaseous carbon dioxide and their affects are rather different. [Twitter, ID: 1305486794360840192]

Conceptualization III: *carbon* is a gas. Scientific evidence of climate change and the dissemination of this information through mass media have contributed to the general public's perception of carbon dioxide (commonly abbreviated as *carbon*) as a primary greenhouse gas. As a result, the conceptualization of *carbon* as an equivalent of *carbon dioxide* is also common in our subcorpus. Additionally, *carbon* is associated with other greenhouse gases. Given the heterogeneity in the specific gases that individuals in our subcorpus associate with *carbon*, we propose further subdivision into more specific conceptualizations of *carbon* as a gas:

C. III.1. $carbon \equiv carbon\ dioxide$

C. III.2. $carbon \not\equiv methane$ ($carbon \equiv carbon\ dioxide$)

C. III.3. $carbon \equiv carbon\ dioxide/methane$

Conceptualization III.1: *carbon* $\equiv$ *carbon dioxide*. As demonstrated in Figure 6.1, the prevalent keywords within our subcorpus are `co2`, `carbon dioxide` and `methane`, with `co2` being obviously dominant. Consequently, we have noticed a tendency in our subcorpus to use *carbon* in a semantic conjunction with *carbon dioxide*. This tendency can be attributed to the common practice of using *carbon* as a shortcut for *carbon dioxide* in environmental context. In texts that represent this concep-

tualization, *carbon* and *carbon dioxide* coexist as contextual and interchangeable synonyms.

For instance, in (11), both *carbon* and CO_2 refer to the greenhouse gas carbon dioxide emitted by Americans and meant for sequestration. In (12), the author discusses the removal of *carbon*/CO_2 through trees and other vegatation. In (13), the author refers to the byproduct of human and vegetation respiration as both *carbon* and CO_2. In (14), the author emphasizes the importance of addressing carbon dioxide emissions, while equating *carbon* with CO_2 that is added into the air.

(11) The average American emits 44 tons of CO2 per year. This facility will sequester the **carbon** of 2.2k Americans, aka 0.0007% of the US population.
[Reddit, ID: enkl5rz]

(12) Are trees the best culture to suck **carbon** out of the air? With recent reports on seaweeds plantations, and permaculture being more efficient than trees at removing CO2, do you have plans to promote to "invest" on those practices?
[Reddit, ID: eweydb6]

(13) We breathe out **carbon** as a waste product of respiration. Plants reverse respiration to store energy, but also respirate themselves. This means plants ALSO breathe out CO2. Plants only absorb more than they emit while they're growing. Burning plants releases all they've absorbed. [Twitter, ID: 1303956853554999296]

(14) I don't think Greta said we can get rid of the **carbon** in the air. We have to adapt, but we also have to stop adding more co2, right now.
[Twitter, ID: 1302917388522971136]

Furthemore, *carbon* is interconnected with *carbon dioxide* through various idioms such as ⌜*carbon emissions*⌝ (15), ⌜*carbon footprint*⌝ (16), ⌜*carbon budget*⌝ (17), and ⌜*carbon neutral*⌝ (18). For example, in (15), *carbon emissions* are linked to the concentration of CO_2 in the atmosphere. In (16), the two idioms ⌜CO_2 *footprint*⌝ and ⌜*carbon footprint*⌝ are used interchangeably. Example (17) illustrates that the increased CO_2 emissions will strain the *carbon budget* by 2024, indicating the point of no return. In (18), the author states that carbon neutrality implies the balance between emitted and removed CO_2, in this specific case by means of trees.

(15) We are in an absolute crisis with levels of CO2 And the government is just not acting anything like fast enough. We need @username[2] to wake you up. Please make radical changes to help us all cut our carbon emissions. Solar panels on every roof, flight tax, planting trees... [Twitter, ID: 1303058381670932480]

(16) But with good leadership, the US could reduce CO2 footprint by eliminating coal fired plants in only a few years. Increase nuclear, efficiency, wind, solar, these take more time, but within 15 years, the US should be able to reduce its **carbon footprint** by 75%. [Reddit, ID: e93vcqz]

(17) Meanwhile, because we haven't cut CO2 emissions since 2018 as the IPCC sr15 dictated but instead expanded them by roughly 3% per year, our **carbon budget** will run out in 2024, meaning that we have 4 years to to cut 100% of our CO2 emissions, which in a best case scenario will still still bring us well above +1.5C, and very likely past +2C. [Reddit, ID: g6v8g4s]

(18) Trees have a limit on what they can absorb and the world as a whole is producing more CO2 so we are not anywhere near **carbon neutral**.
[Twitter, ID: 130268660170627072]

The semantic relation between *carbon* and *carbon dioxide* can also be traced through the choice of measurement units used to estimate the environmental impact of carbon dioxide emissions. Example (19) discusses the concept of carbon tax as a financial mechanism to address excessive carbon dioxide emissions by proposing financial charges ranging for $10 to $50 based on tons of the emitted CO_2.

(19) No, all we need is a **carbon tax** of between $10 to $50 per ton of CO2. The money from that tax then ought to go directly to **carbon offsets**. Tree planting, biochar, enhanced wethering etc. That way we could be **carbon negative** (globally) in a few years. The time is only dependant on the time it takes to scale the offsets. [Reddit, ID fbj9x9x]

In conclusion, conceptualization of *carbon* as a synonym of *carbon dioxide* is prevalent in our subcorpus and it corresponds to the common practice of using the Term *carbon* in the official environmental discourse and press. Nevertheless, the coexistence of *carbon* 'chemical element' and *carbon* 'carbon dioxide' in environmental discourse appears to cause confusion among the general public. For

[2]We anonimize the real user names in mentions according to the Twitter Terms of Service.

example, consider a conversation snippet between two users in (20). User A questions the statement that carbon is bad for the planet and contributes to global warming. User B points out that user A's reasoning may be flawed due the confusion between carbon as a constituent of carbon compounds and carbon dioxide as a greenhouse gas. We believe that this dual usage deserves specific attention from educators.

(20) A: Tell me how **carbon** is a bad thing for the planet. Or how the warming started before **carbon** increased? I'm all for reducing 'pollution' totally with you on that (though not if you're ignoring China and India), but not '**carbon**.' [Twitter, ID: 1302659794331930626]

B: You're confusing **Carbon** which is found in many forms and is important to life and carbon dioxide which until recently was only present in small amounts in the atmosphere. Our recent activities have increased Co2 to levels which are causing climate change. [Twitter, ID: 1302703778735501312]

Conceptualization III.2: *carbon* $\not\equiv$ *methane* (*carbon* $\equiv$ *carbon dioxide*). This conceptualization is quite similar to the preceding one, with an important semantic nuance. While *carbon* is still used synonymously to *carbon dioxide*, there is an additional explicit dissociation with *methane*. In (21) and (22), the authors compare the heat-trapping potency of methane to that of carbon, i.e., carbon dioxide. Although the latter example does not explicitly mention carbon dioxide, we can still assume that methane is being compared to carbon dioxide. This comparison between methane and carbon/carbon dioxide illustrates the perception of these as different concepts.

(21) I'm certainly not an atmospheric chemist, but even though methane is a more potent greenhouse gas than CO_2, it dissipates incredibly quickly in the atmosphere when compared to **carbon**. [Reddit, ID: d3y89p0]

(22) Record Arctic temps triggering giant methane explosions. Methane traps 30x more heat in the atmosphere than **carbon**. @username @username lets get going on decarbonization! [Twitter, ID: 1301198240554135557]

Such perception of *carbon* is also applicable to idioms. In (23) and (24),

carbon emissions are distinguished from *methane emissions*. Furthermore, in (25), *methane* and *methane emissions* are contrasted with ⌜*carbon footprint*⌝ and ⌜*carbon dioxide*⌝.

(23) Could be something I don't know, but this article is talking about methane emissions. Doesn't seem to say anything about **carbon emissions**. [Twitter, ID: 1302217591373660161]

(24) Hi Tesco, where does the soya that feeds the animals you sell as meat come from? Since you care about deforestation, have you looked at your supply chain? And maybe it's time to sell less meat? Would help reduce methane and **carbon emissions** a bit... Just a thought.... [Twitter, ID: 1303055296131145735]

(25) It's not about the **carbon footprint**. It's about methane. Cattle are a leading source of methane emissions in the US which is a worse greenhouse gas than carbon dioxide. [Reddit, ID: emrc15o]

Conceptualization III.3: *carbon* ≡ *carbon dioxide/methane*. Interestingly, this conceptualization is in terminological conflict with the previous one, since in this case the concept of carbon includes methane. Specifically, *carbon* serves as an umbrella expression which includes both carbon dioxide and methane. In some instances, other gases might also be included, although the authors do not always specify which ones. This conceptualization of carbon as a greenhouse gas corresponds to the sense CARBON **II.1** (see 5.5.2.3).

In (26), the author questions someone's understanding of carbon emissions, while explicitly stating that, from their viewpoint, carbon emissions include not only carbon dioxide but also methane emissions. Example (27) is a good illustration of *carbon* as an umbrella Term whose reference can only partially be inferred from the context. The first part of the text explicitly indicates that *methane* is included in the meaning of *carbon* or, more precisely, *carbon footprint*. The second sentence suggests that, apart from methane emissions, there is an additional source of harmful emissions involved in cattle management that contributes to the cattle's carbon footprint. This source of emissions and the greenhouse gas itself are not directly mentioned and cannot be inferred from the text. Similarly, in (28), the author discusses the carbon cost of livestock management and beef production in terms of related methane emissions, while also pointing out that

non-methane emissions should be considered as well. In (29), the carbon footprint of meat consumption is associated with methane emissions from cattle raising.

(26) You realize when they say **carbon emissions** that doesn't just mean CO_2 right? Methane is far more problematic. [Twitter, ID: 1301372101039976448]

(27) The **carbon footprint** of cattle is only slightly due to their methane belching. Much more is due to the inputs required to raise the animals. [Reddit, ID: epyy0uf]

(28) There's much more to livestock-related **carbon costs** than just methane. As highlighted below, most of beef's production emissions come from non-methane sources. And most of beef's **carbon costs** don't even come from production emissions. [Twitter, ID: 1301304644107370496]

(29) Our diets are a huge **carbon footprint**. If you measure the effects of **CH4** over a 10 year period, then omni diets have a huge footprint. Add in the deforestation associated with raising beef, the insane amounts of **methane** in dairies - and here's the thing - it's all unnecessary! [Reddit, ID: eimnder]

We have also discovered instances of more implicit semantic connection between carbon and methane. In (30), the author talks about carbon, specifically in the context of carbon pricing, as a mechanism to address CO2-e contributors. *CO2-e* stands for carbon dioxide equivalent, a measure comparing the environmental impact of greenhouse gases to that of carbon dioxide. Given the common association of cattle farming with methane emissions, we can assume that CO2-e includes methane. However, potentially, it could include any greenhouse gas, as CO2-e measure was specifically developed to account for any greenhouse gas emissions in relation to the predominant greenhouse gas carbon dioxide.

(30) The most impactful action available is a price on **carbon**, which systemically addresses all CO2-e contributors, including agriculture. More importantly, a price on **carbon** actually moves people towards becoming vegetarian, as opposed to just telling people. A **carbon price** causes the price reflect the true cost of unsustainably-farmed meat. [Reddit, ID: erbq8fw]

6.3 Further insights

While exploring the corpora beyond the scope of the compiled subcorpus, we detected interesting linguistic insights related to *carbon* that complemented the

objectives stated in the preamble of the chapter. These insights pertain to the linguistic behavior of *carbon* in Twitter and Reddit texts and language peculiarities of social network corpora. Specifically, we will discuss an interesting blend of specialized and colloquial language, abstracted perception of carbon, terminological debates, and ideological framing of carbon-related discussions.

6.3.1 Colloquialization

During the process of domestication, a Term often becomes a part of the general language as a runaway Term or a quasi-Term. As a consequence, it adapts to a more colloquial and less specialized lexical environment. This means that the specialized lexical units that the Term previously interacted with in the specialized discourse are often replaced by their more colloquial counterparts in everyday language.

Specifically, we have noticed that users tend to describe industrial and natural processes associated with *carbon* in a "simplistic" and metaphoric manner. This is particularly noticeable in the way verbs are used to denote these processes: *pump carbon into the air* (31), *pull carbon out of the air* (32), *suck carbon* (33), *dig up carbon* (34), *put carbon into the air* (35), and *cancel out the carbon* (36).

(31) Simple things like turning off light switches when you leave the room, not running the faucet while brushing your teeth... when you multiply these kinds of tiny things across millions of people it translates into tens of millions of litres of water saved, or who knows how many gigatonnes of **carbon not pumped into the air** by a power plant. [Reddit, ID: d5trqhu]

(32) Just **pulling carbon out of the air** is possible, but it requires a whole lot of energy. [Reddit, ID: eh91wxg]

(33) Trees take 100 years to fully grow. The time for trees was a long time ago. Whatever we do, it has to **suck carbon** faster than that. [Reddit, ID: f6er2qg]

(34) We have to be willing to **dig up less carbon**. Unless we put some limits on **digging up carbon**, we're not going to solve this. [Reddit, ID: ees3lt6]

(35) You obviously don't realize the 100s of tons of **carbon you are putting into the air** with your holier-than-thou attitude of "I have a job so I can afford to commute" attitude. [Reddit, ID: eesc794]

(36) I'm not saying planting trees is bad. I am just saying that it shouldn't be looked at as a way to **cancel out the carbon** you are emitting. [Reddit, ID: f6k9sok]

These linguistic choices create an interesting blend of scientific terminology and colloquial language, demonstrating how the scientific buzzword *carbon* coexists with colloquialisms. Table 6.1 provides a list of scientific equivalents of some of the simplified verbs used in the above-mentioned sentences and other texts from our corpus.

Colloquial expressions	Scientific equivalents
carbon not pumped into the air by a power plant, *pull carbon out of the sky and put it in the ground*	*carbon sequestration*, *soil carbon sequestration*
pull/suck/remove carbon out of the air	*carbon capture from the atmosphere*
dig up carbon, *pump carbon out of the ground*	*carbon mining*, *fossil fuel extraction*
cancel out the carbon	*offset the carbon emissions*
put carbon into the air	*emit carbon into the atmosphere*
atmospheric carbon dumping	*excessive carbon emissions*
burn carbon	*burn fossil fuels*, *fossil fuel combustion*
carbon dioxide above our heads	*atmospheric carbon dioxide*
grass absorbs carbon from the air and soil	*photosynthetic carbon (dioxide) uptake*
type of concrete that instead of making carbon it sucks it in and locks it away	*carbon-negative concrete*
trees soak up the atmospheric carbon	*atmospheric carbon absorption*

Table 6.1 – Colloquial expressions concerning *carbon* and their scientific equivalents.

The colloquialization can also be intentionally used as a discourse technique with the purpose of making scientific knowledge more accessible for non-experts. In a rather peculiar example (37), the user addresses the topic of carbon cycle within a living tree. They use a blend of specialized Terms (such as *carbon chain*, *organic acid*, *humic acid*, *fulvic acid*) and colloquial expressions. Interestingly, the user utilizes the Term *exudate*, even though it is misspelled (as *edudates*)

twice. We find simplified and metaphoric verbs concerning carbon, such as *taking carbon out of the air* and *putting it into* [*tree's*] *body*. Furthermore, informal language in this post includes such expressions as *pee* and *poo* in reference to the processes and byproducts of microorganisms' vital activities. The respective scientific equivalent would be *microorganisms excrete waste products*. This excerpt demonstrates that the author uses their expertise to clarify some scientific details for their interlocutor, yet the seeming expertise on online forums can sometimes be misleading. For people seeking answers, it would be a better practice to consult scientific literature on the subject for fact-checking.

(37) When the tree is alive, it is not only **taking carbon out of the air to put it into it's body** (literally growing out of thin air), it is actually also secreting these **carbon chains** called **plant-root-edudates** into the soil. Often this is called "**liquid carbon**". Think of these **edudates** like little bacteria and fungal treats. The soil microbiology eats these treats and "sequester" the carbon and nutrients into their body. Now these microorganisms **pee** and **poo** and die. These 3 processes bind, or "chelate" the nutrients onto **organic acids** like **humic acid** and **fulvic acid**." [Reddit, ID: f2vz885]

In another interesting example of specialized language colloquialization, (38), the user mocks their interlocutor for using "fancy" terminology such as *photosynthesis*, suggesting that the same concept could be expressed in a less sophisticated way by simply referring to the act of planting trees.

(38) That's a lot of fancy sounding words to say very little. Ooh, you're using the "ability to capture carbon dioxide through **P H O T O S Y N T H E S I S**"? How high tech! So like, planting trees, then. [Reddit, ID: fd9uu17]

6.3.2 Discourse of personal responsibility

With the emergence of the concept of personal carbon footprint (see 5.5.5), the responsibility for negative environmental impact has shifted significantly towards the general public, at a personalized level. One of the most conspicuous examples of such a shift is the promotion of accessible carbon footprint calculators. Furthermore, due to various financial solutions for environmental mitigation, carbon has become a marketable commodity. This was reinforced by the development of carbon tax, personal carbon allowances, etc. Such financial mechanisms allow for

the taxation of large organisms such as companies as well as individuals by introducing charges for greenhouse gas emissions. For example, the system of personal carbon allowances implies the allocation of budget of carbon units to adults, which they can use to pay for transport, home-heating fuels, and electricity (Fuso Nerini *et al.* 2021).

Conceptually, these approaches to managing carbon emissions have significantly contributed to how one perceives carbon. In mitigation context, carbon has transformed into carbon units, which are "complex abstractions that exist only by way of agreement" (Paterson and Stripple 2012). This perception is reflected in the language, specifically in the environmental discourse related to excessive carbon emissions and ways to reduce them. In our corpus analysis, we have noticed a clear trend towards the personalization of carbon pollution, linguistically evident in at least two ways: (i) the use of personal possessive pronouns such as *my*, *your*, *his*; (ii) the use of verbs that take an individual as a subject, such as (*use*, *consume*, *emit*, *save*).

In (39), the author call themselves a $CO_2(e)$ *emitter*, implying their relatively sustainable lifestyle compared to others. Furthermore, they also mention *personal CCS methods* to counteract the previous actions which contributed to emissions, although CCS (short for *carbon capture and storage*) is an industrial technology primarily aimed at large-scale processes. The verb *emit*, commonly associated with countries, factories, and processes, is used here to describe personal emissions. Similarly, in the following examples, the verb *emit* is applied to vegans (40) and Americans (41). Additionally, in (41), the rich individuals are accused of their *carbon use*, which is presumably synonymous to *carbon emissions*. Interestingly, in (42), the verb *emit* is applied to a latte drink, implying that the production of its ingredient, specifically milk, is linked to excessive methane emissions and contribute to overall carbon emissions.

(39) Yet I'm not the one spreading literally evil denier crap across the internet. And I'm fairly sure I'm in the bottom 1-2% of **CO2(e) emitters** in my country. I'm doing my share, *and* I'm interested in **personal CCS methods**, and the developments of such, so I can "get rid" of the stuff I emitted previously and, of course, actually try and *save humanity*. [Reddit, ID: em074y6]

(40) You're still emitting **carbon** on a plant based diet, along with all the same industrial mechanisms designed to put it on your supermarket shelf. Think. Stop it with just the vegan propaganda. [Reddit, ID: e7ks04j]

(41) The most effective thing we could do to immediately address household consumption would be to ration the **carbon use of the rich**. The top 10% of Americans emit 50 tons per person per year due to their consumption. The bottom 50% emit about 10 tons per person per year. [Reddit, ID: epxwbok]

(42) This sort of figure makes me glad I'm a black coffee drinker. One latte emits the same as almost a fifth of an average person's daily food related **carbon emission**! All that milk from methane burping cows. [Twitter, ID: 1301413121077239808]

The comparative analysis of the use of the verb *emit* in the above-mentioned examples with its use in specialized environmental discourse reveals a subtle semantic difference. In specialized discourse, the phrase *power plant emits carbon* denotes the fact that a power plant releases greenhouse gases as a result of its operations, **directly** contributing to negative environmental impact. In contrast, the phrase *the rich emit carbon* refers to individuals' actions that **indirectly** contribute to negative environmental impact. Indirect contribution involves such activities as driving cars, flying, consuming meat, and making unsustainable consumer choices.

Furthermore, notions originally applied to sustainable industries and technologies are now being applied to individuals, such as *carbon negative* (43) and *zero carbon* (44). Zero carbon technology refers to technologies, industries, processes with minimal carbon emissions or no emissions. In contrast, the notion of a zero carbon person shifts the focus to the individual level, referring to individuals who adopt sustainable practices in daily life to reduce their personal carbon footprint. While the concepts of zero carbon technologies and zero carbon individuals both describe approaches to environmental mitigation, they entail different perspectives and mitigation tools.

(43) He could be the first milionair (or however much he owns) to be a **carbon negative** person while living a good life. [Reddit, ID: f62vo0w]

(44) Already working on becoming **0 carbon** with the wife it's hard as heck. but what's the alternative. [Reddit, ID: eat7a32]

These contexts have led us to consider the existence of a new sense of *carbon* pertinent to the general language. As a result, we have introduced a separate sense CARBON **II.2** within CARBON vocable in the English Lexical Network (see 5.5.2.4). This sense accounts for the above-mentioned metaphorical usages and similar cases, where speakers do not directly refer to greenhouse gases but rather to everyday activities associated with their emissions.

6.3.3 Confusion and debates about terminology

As mentioned in 5.3, the fast evolution of carbon-related terminology often leads to confusion among the general public as they struggle to understand the meaning of newly introduced Terms. Our corpus has provided evidence that supports this observation. Specifically, we have identified two main categories of discussions. The first category deals with explicit confusion about environmental terminology. In the second, we find texts where authors critique others for their scientific reasoning and choice of terminology. The majority of such debates revolves around topics such as the carbon cycle, the meaning of carbon idioms (⌜*carbon neutral*⌝, ⌜*carbon free*⌝, ⌜*zero carbon*⌝), and climate change skepticism.

In the discourse of confusion, users express uncertainty about the meaning of certain expressions, which makes them inquire additional explanations. For example, in (45) and (46), the authors express doubt whether carbon tax covers other greenhouse gases such as methane. These examples illustrate how the ambiguity of *carbon* (discussed in 5.4.1 and 6.2) confuses the general public making them question whether *carbon* exclusively refers to carbon dioxide or serves as an umbrella expression. In (47), the author asks for clarification of regional differences concerning the notion of *carbon-free* as it may not refer to the same concept in the United States and Europe. Such texts emphasize the need of clear and unambiguous terminology as a way to minimize confusion.

(45) When people say **carbon tax** do they ever mean it as a shorthand for a broader GHG tax? We need to tax **methane** to affect meat consumption, right? [Reddit, ID: ecj3wbn]

(46) Is it **carbon intensive** or **methane intensive** or both? Do any of the **carbon taxes** also tax **methane**? [Reddit, ID: ewpjhgg]

(47) What does "**Carbon-free**" mean in the US? It's not a term I've heard before. Can someone bring insight, a definition? It's confusing to me since in Europe wind and solar are not seen as being having no emissions, it's seen as "low emission" as they still have emissions and waste associated with the manufacturing process, maintenance and the recycling process at EOL (some panels for instance contain toxic waste like cadmium and lead so they need proper recycling). [Reddit, ID:gkmlvym]

Furthermore, another cluster of texts contains discussions where users correct and comment on each other's way of reasoning and on expressions, often using aggressive and explicit language. Example (48) shows a snippet of a conversation between users B and C who discuss the term *carbon neutral* after user A states that Formula 1 driver Lewis Hamilton offsets his carbon footprint both in personal and professional life. User B questions Hamilton's carbon neutrality and assumes it might be a PR strategy. As user B goes in details of what they understand by being carbon neutral, user C concludes that certain statements are uninformed and that user B does not understand the underlying science of carbon neutrality and carbon offsetting. Specifically, user C points out that being carbon neutral does not imply that humans have to stop breathing, formally considered as a source of CO_2. This conversation highlights the different intepretations of the term *carbon neutral* that individuals might have. For example, user B's statements show they have quite exaggerated perception of carbon neutrality which does not correspond to its scientific use.

(48) A: lewis offsets his carbon footprint both on the track and off, he is **carbon neutral** right now. [Twitter, ID: 1303410972393181184]

B: You can't actually believe what you've posted? He's **carbon neutral**? Does he drink rain water and shit in a composting bin? No, he uses public utilities like you and I that run off fossil fuels. He uses plastic (fossil fuel).... etc etc etc. [Twitter, ID: 1303433309205204994]

C: You absolutely don't know what a carbon footprint is, and you don't know what the word „offset" means [Twitter, ID: 1303548347421253632]

B: You really think this puppet is really offsetting his carbon footprint? It's all

> PR bs that they sell to us. Only some of us are really that naive to believe it though... [Twitter, ID: 1303652231325970432]

> C: No, what I think is he pays some organization a bunch of money which plants some trees and claim to offset a specific amount of CO2, and does that only for PR reasons while absolutely not caring. But it's not like you said that it's impossible for him to be **carbon neutral**. [Twitter, ID: 1303658350517813253]

> B: Does he also live in the forest? Travel by foot? If not, is his tree farm harvested by hand (not tools because those are literally made from steel with lots of carbon). If you live like a wild animal, then yes I'll believe you can offset your carbon foot print. [Twitter, ID: 1303659298715705347]

> C: Okay I was right - you absolutely don't know what it means to live **carbon neutral**. It doesn't mean you have to stop breathing because that also produces CO2, but the equation at the end has to be a 0. Also you wouldn't harvest a tree farm if your goal is to offset CO2 [Twitter, ID: 1304130795305603072]

6.3.4 Ideological framing

In environmental discourse, we observe polarized views on the reduction of carbon emissions. On the one hand, carbon dioxide is often portrayed as the main culprit of climate change, and the ultimate goal of environmental mitigation is seen as drastically reducing or completely eradicating any man-made carbon dioxide emissions. As a consequence, carbon dioxide, and therefore carbon, is generally perceived as a harmful pollutant. On the other hand, experts warn that it is factually inappropriate to reduce this debate to the fact that carbon dioxide is inherently harmful. Specifically, they point out that "demonization of carbon" is ureasonable and it compomises its crucial role in fundamental processes such as the carbon cycle:

> Without carbon and carbon dioxide, there would be no life on earth. [...] The conclusion that carbon dioxide is a pollutant has no scientific foundation, and when carbon itself is regarded as a 'pollutant' the situation is utterly non-scientific (R. Evans 2009).

In our corpus, we identified a collection of texts where the authors support the latter perspective. These authors emphasize the significance of carbon and protest against its unjust stigmatization. In particular, the authors reject concepts that imply the limitation or complete elimination of carbon dioxide or carbon altogether, such as *zero carbon* (49), *carbon neutrality* (50), and *carbon capture* (51). One of their main arguments is that carbon dioxide serves as plant food and that life on Earth is *carbon-based*, as in (49), (52), and (53).

The one-sided views expressed by these users partly stem from the terminological problem discussed in 6.2. This concerns the duality of *carbon* when it comes to two contextually and semantically related senses 'chemical element C' and 'carbon dioxide'. In the given examples, the notions of carbon footprint, zero carbon, carbon neutrality and carbon tax are considered with chemical element carbon in mind, which contributes to an erroneous understanding of their meaning. Although these notions do refer to practices aimed at the reduction of emissions, none of them implies the complete eradication of existing carbon dioxide, and even more so of the chemical element carbon, on the planet as a whole. For example, *zero carbon* means that "no carbon emissions are being produced from a product or service."[3] This practice aims to eliminate the excessive carbon dioxide emissions associated with industrial production, not the carbon dioxide that naturally exists as a component of the atmosphere.

(49) Plants need **carbon**, we are **carbon based lifeforms**. how do you get to **zero carbon**? This is another way to try and push a depopulation agenda. And the real fight of our generation is human trafficking and slavery world wide. Not "global warming" that's just a scam, Not buyin it. [Twitter, ID: 1300516459328745472]

(50) **Carbon neutrality** is a lie. All life depends on **carbon dioxide**. If life was "**carbon neutral**", you'd no longer exist. Plant life depends on **CO2** for food. Stop the hoax - fossil fuels are not the evil it is reported to be.
[Twitter, ID: 1303428225310306305]

(51) '**Carbon capture**'? What in God's name are you talking about? I take it you're talking about **carbon dioxide**, that non-toxic, non-pollutant, trace gas that is essential for life on this planet. [Twitter, ID: 1303476496217104388]

[3]https://www.nationalgrideso.com/future-energy/our-progress-towards-net-zero/net-zero-explained/what-net-zero-and-zero-carbon

(52) As everything on this planet is **carbon-based**, it is impossible to NOT have a **carbon footprint**. **CO2** is not the same as **carbon**, BTW. They use both interchangeably, though. Which is weird. The only way to have a **zero-carbon footprint** is to cease to exist. [Twitter, ID: 1303121444562284544]

(53) I propose stopping calling it a **carbon tax**, vilifying the innocent element, and calling it instead an "**externality tax**," "**pollution tax**," or even "**overheating the Earth tax**" if you're not into the whole brevity thing. [...] I support shifting taxes from some areas to taxing greenhouse gas emissions, but as far as language goes, as a **carbon-based life-form**, I don't like the sound of taxing the basis of my life. [Reddit, ID: gc5y67z]

6.4 Discussion

After manually analyzing 3,204 tweets and 1,591 Reddit posts (1,552 comments and 39 submissions), we gained valuable insights into how Twitter and Reddit users engage in discussions related to the environment. Specifically, we discovered how communication on these social networks functions, the differences and similarities between them, and more importantly, how users express their views on carbon-related environmental issues.

We found that in both social networks, the users do post environment-related content and participate in environment-related conversations, exchanging their knowledge, experience and opinions. Our subcorpus was limited to carbon-related issues, hence the texts primarily discussed carbon dioxide emissions, low carbon technologies, the carbon cycle, carbon footprint, carbon markets, carbon neutral future, etc. Within these discussions, we identified three main conceptualizations of *carbon*: as a chemical element, a solid, and a gas, with further subdivision within the gas conceptualization.

While we did not detect any distinct conceptualizations unique to either tweets or Reddit posts, we acknowledge the differences between the two corpora. These differences, which include the length of texts and communication structure, originate from distinct nature of each platform (see 4.2). As previously mentioned, the average length of tweets in the subcorpus was 38 tokens, whereas Reddit posts were 218 tokens long. Extensive Reddit posts allowed us to study the occurrences of *carbon* in more rich and meaningful context. Furthermore, our analysis confirmed that tweets represented rather isolated texts, while Reddit comments often

formed comprehensive conversation threads.

Considering these points, we concluded that Reddit is more suitable for qualitative linguistic analysis due to its lengthier and more substantial texts compared to tweets. Nevertheless, Twitter remains a valuable source of the general public interactions. While some academic projects have reported limitations of working with Twitter data (for example, O'Connor *et al.* (2014) mentioned the scarcity of medicine-related Twitter data in comparison to health forums), we did not encounter such a problem. The primary limitation we encountered while working with Twitter data was the lack of context and brevity due to the character limit. However, this limitation did not ultimately prevent us from obtaining meaningful insights about the use of *carbon* in tweets.

Regarding the language use on these platforms, we have noticed that due to their informal nature and absense of strict communication rules, users feel free to use explicit, aggressive, and politically incorrect language. At the same time, such informality does not discourage them from using scientific terminology. By filtering our main Twitter and Reddit corpora with specific keywords to compile the subcorpus, we obtained data where users express their opinion and engage in conversations using scientific terminology related to carbon emissions. Clearly, not all instances of terminology usage conform to scientific standards, as the analysis of conceptualizations of *carbon* revealed. Furthermore, what was more interesting to us, the combination of the platforms' informality and environmental context of discussions created an interesting blend of colloquial and scientific language.

Concluding our findings of the use of *carbon* in the specialized discourse (see 5.4) and the ordinary discourse, we have formulated a list of recommendations to improve environmental terminology and its dissemination. These recommendations also extend to communicating scientific aspects related to carbon emissions.

1. It is necessary to revise the current use of *carbon* in the specialized discourse in order to find a more clear and consistent way of addressing the issue of greenhouse gas emissions. For example, certain experts propose using the Term *greenhouse gas* as a more scientifically appropriate alternative, particularly for idioms, e.g., *greenhouse gas emissions*.

2. We suggest putting a specific emphasis on the distinction between *carbon* 'chemical element' and *carbon* 'carbon dioxide' to avoid biased views among the general public. This differentiation is crucial for understanding environmental solutions aimed at reducing excessive greenhouse gas emissions rather than eliminating the chemical element carbon itself.
3. One further suggestion would be to reconsider the use of *carbon* as an overarching expression in idioms that refer to greenhouse gas emissions and their reduction, e.g., ⌜*carbon neutral*⌝, ⌜*low carbon*⌝. First, the focus on *carbon* leads to one-sided perception which neglects other greenhouse gases. Second, the similarity in such idioms can lead individuals to mistakenly believe that they refer to the same concept of carbon, which is scientifically inaccurate. For example in 5.5, we demonstrated that *carbon* in idioms ⌜*carbon cycle*⌝, ⌜*carbon footprint*⌝ and ⌜*carbon capture and storage*⌝ refers to three different concepts.
4. Building on point 3, it is necessary to provide clear and consistent definitions for idioms used to denote emission reduction practices, such as ⌜*carbon capture*⌝,⌜*carbon free*⌝, ⌜*carbon neutrality*⌝, ⌜*carbon removal*⌝, ⌜*low carbon*⌝, and ⌜*zero carbon*⌝, etc.
5. To address the scientific aspects, we recommend to clarify the difference between two distinct concepts associated with carbon dioxide. On the one hand, carbon dioxide is a natural constituent of the Earth's atmosphere along with nitrogen, oxygen and other gases, and is also produced through human respiration. On the other hand, elevated concentrations of carbon dioxide in the atmosphere contribute to climate change and impact the Earth's termperatures. These interpretations of carbon dioxide as a vital component and a harmful greenhouse gas often get mistakenly conflated.
6. Finally, we recommend educating the general public about greenhouse gases beyond carbon dioxide. At present, greenhouse gas emissions are primarily associated with carbon dioxide (due to the proliferation of *carbon*) among the general public. However, scientific insights on climate change entail more than just carbon dioxide.

These recommendations are addressed to lexicographers, terminologists, ex-

perts, educators and activists, and presuppose their collective efforts. The recommendations concerning terminological aspects was partly addressed in our lexicographic work described in 5.5. Specifically, we uncovered the polysemy of the vocable CARBON taking into account its chemistry, environment and general language senses. Furthermore, we worked on the related idioms. Clearly, future research should consider other lexical units that belong to *carbon*-related terminology.

6.5 Chapter synthesis

We conducted a corpus analysis to explore the perceptions of *carbon* when used in ordinary discourse. For this, we built a "*carbon* subcorpus" using specific keywords and performed a qualitative analysis of extracted texts with three main objectives: (i) to assess the potential of using social network texts for linguistic analysis; (ii) to study the conceptualization of *carbon* in ordinary discourse; (iii) to identify the semantic links between *carbon* and other related environmental vocabulary.

The findings revealed several senses of *carbon* in ordinary discourse: carbon is commonly conceptualized as a chemical element, a solid and a gas. Furthermore, we identified the use of *carbon* that refers to 'symbolic polluting substance'. The sense was included in the English Lexical Network as CARBON **II.2**. Through an extended analysis of texts with *carbon*, we discovered various clusters of texts characterized by colloquialization, abstracted perception of carbon, terminological debates, and ideological framing.

Finally, we demonstrated conflicting usages of *carbon* which emphasize the need for a review of *carbon*-related terminology and its further normalization at both national and international levels. To address this challenge, we formulated a list of terminological recommendations to improve environmental communication.

Chapter 7

General conclusion

Summary

7.1 Summary of findings

In Chapter 2, we began our study with an attempt to define the place of the environmental domain within the international agenda by considering the broad environmental context that is rooted in sociology. Through the examination of interdisciplinary aspects of environmental science and public engagement with the subject, we identified common points of interactions. Such interactions include global environmental awareness, collaborative public efforts in mitigation, expert-layman environmental communication, public uncertainty and mistrust, and environmental literacy. We concluded that environmental domain in its current state has significant social implications and extends beyond academic concerns.

In Chapter 3, we analyzed distinct linguistic features of environmental discourse, which is characterized by polyphony of voices, dynamic nature, terminological variation, and interdisciplinarity. To narrow our focus, we opted to concentrate on the terminology of current and emerging environmental issues. We applied a hybrid approach, which involved NLP tools and manual selection, to compile a list of 290 keywords (Chapter 4). These keywords serve as proxies for environmental Terms that shape the topic of current and emerging environmental issues.

As mentioned earlier, we do not consider the list of keywords exhaustive and, quite on the contrary, the aim for it is to be constatly updated and further expanded. There are two main ways of expanding the list: (i) keeping track of emerging environmental concepts that become relevant to environmental domain; and (ii) converting the existing keywords into meaningful lexical units for lexicographic description in a lexical network. We have already formed a foundation for the second point by describing *carbon* and the related idioms in the English Lexical Network (Chapter 5). By accounting for their combinatorial properties, we connected them to other lexical units such as *methane*, *greenhouse gas*, *atmosphere*, *climate*, etc. On the one hand, due to such systematic approach, it will be possible to connect all the lexical units from our keyword list which will result in a large comprehensive lexical network of interconnected environmental Terms. On the other hand, it will help to reach "new" lexical units relevant for environmental topic.

Furthermore, the applications of the proposed keyword list extend to other tasks and purposes:

- compilation of thematic environmental glossaries;
- development of educational materials and courses;
- NLP applications;
- translation purposes.

Using the identified keywords, we created two social network corpora from Twitter and Reddit (Chapter 4). To perform the data extraction, we analyzed their architecture and defined distinctive extraction techniques for each social network. The extracted data was further used for a qualitative analysis. We found that in both social networks, users do post content on the environment and participate in environmental debate, while exchanging their knowledge, experience and opinions. We concluded that Reddit is a more adequate source for meaningful qualitative analysis, whereas Twitter is valuable for studying environmental lexicon in a very consice and restricted context, which also affects how people communicate.

To account for carbon-driven environmental agenda, we chose *carbon* as an object of our linguistic study. In our comprehensive analysis of *carbon* in Chapter 5, we first addressed the sociological aspects of its place in current environmental talks. We found that numerous sociological, political, economical, technological, marketing, and layman concepts are related to *carbon*: carbon economy, carbon neutral society, low carbon future, carbon colonialism, carbon detox, carbon footprint, etc. In our linguistic analysis, we observed the complication of *carbon*-related terminology, acknowledged by both specialists and the general public. We indentified that this complication is rooted in the contextual fluctuation of meaning of *carbon*, interchangeable use of synonymous expressions, and lack of description in lexicographic and educational sources.

During the following stage described in Chapter 6, we performed a qualitative analysis of *carbon* in social network corpora. We found three distinct conceptualizations of *carbon* as a chemical element, a solid and a gas. Furthermore, complementary analysis showed that users employ a blend of colloquial and scientific language, discuss personal responsibility in emission reductions, occasionally en-

gage in debates about environmental terminology, and use ideologically connotated expressions. To sum up, both the analysis of specialized and ordinary discourse revealed that *carbon* is primarily associated with *carbon dioxide* and is used as a shorthand expression. These findings formed the basis for our lexicographic analysis aimed at developing the vocables in the English Lexical Network.

The theoretical and descriptive framework of Explanatory and Combinatoiral Lexicology proved to be a reliable and effective methodology for lexicographic treatment of terminology. Specifically, we successfully integrated terminological lexical units into the general language lexical system of the English Lexical Network. We focused on the polysemy of *carbon* and ⌜*carbon dioxide*⌝. Additionally, we selected three idioms – ⌜*carbon cycle*⌝, ⌜*carbon footprint*⌝ and ⌜*carbon capture and storage*⌝ – to illustrate that while *carbon* was a common element in these idioms, it refers to three different concepts. In our lexicographic description, we benefited from the collaboration with a specialist in green chemistry and gained valuable scientific expertise in chemistry and environmental domain. Through multiple iterations and meticulous formulation of definitions, our lexicographic articles provide reliable and accurate domain-specific information. To our knowledge, no prior research addressed the language of social networks from the perspective of lexical-semantic analysis, specifically Explanatory and Combinatorial Lexicology.

Within the context of domestication of Terms, we introduced novel typology of Terms based on their semantics and usage in specific register. Furthermore, we highlighted the difference between quasi-Terms and metaphoric senses. We applied this theory in our examination of the vocables described in the English Lexical Network, identifying instances of all three types of Terms. Particularly, we were interested to detect quasi-Terms, which serve as emblematic examples of scientific language adopted by the general public. In this regard, we identified the sense CARBON **II.2** that is common in the general language and refers to 'symbolic polluting substance', or in other words, nonmaterial by-product of one's unsustainable lifestyle. We integrated this domesticated sense into the terminological vocable CARBON, which aligns with recommendations on how to treat determinologized senses in lexicographic sources outlined in Meyer, Varantola and Mackintosh (1998). Such approach contributes to better understanding of domesticated Terms

and their placement within specialized vocables.

Finally, while working on this project, we formed general thoughts on the future of environmental terminology. Clearly, the state of the environment and role of humanity in its preservation will be among top concerns of the future science. Consequently, environmental terminology will continue to evolve and expand. On the one hand, this process will account for new emerging scientific concepts. On the other hand, the general language will continue to produce creative layman expressions. Additionally, with the pressing issue of carbon emissions, scientific and social senses of *carbon* will continue to evolve as well. Without proper management, it will further contribute to terminological complication and variation.

The reliance on the public as an important actor of environmental mitigation makes the effective dialogue between environmental experts and the general public extremely important. A potential strategy to improve this dialogue can be twofold. First, it is necessary to develop a standardized terminology in collaboration between terminologists and environmental experts. Although certain experts advocate for changes in terminology, this issue has not yet gained considerable attention. Second, it is necessary to promote clear and effective dissemination of environmental information. While current environmental communication often targets emotions, it is important to ensure that the public understanding of environmental Terms conforms to scientific standards.

Due to the insights we gained from our linguistic profiling of *carbon*, we have formulated concrete terminological recommendations to improve environmental communication. The key takeaways include: (i) revising the usage of *carbon* as a standalone shorthand expression and within idioms, and (ii) drawing a distinction between *carbon* and *carbon dioxide* in the context of emissions reduction. The latter point is particularly important in the paradigm of decarbonization.

While the consequences of environmental terminology misuse might not be as serious as in medical context, where unclear communication can lead to fatal outcomes, environmentally knowledgeable citizens will be able to make informed decisions, critically assess mitigation initiatives, and confront greenwashing practices.

7.2 Future research

We identify four main areas for future research to build on our work. First, it would be necessary to continue the lexicographic description of environmental lexical units in the English Lexical Network, as mentioned earlier. The compiled keyword list will serve as the lexical material.

Second, we would be interested to extend our analysis to other environmental Terms by applying the same methodology used to analyze *carbon*. Many environmental Terms deserve attention from the perspective of their use within general public discourse. Previous research, including the referenced papers, has already highlighted a number of environmental Terms that cause confusion among the general public. We believe that a logical extension of our research would involve the analysis of the environmental Terms *greenhouse gas* and *greenhouse effect*. Through our work, we have gathered certain insights into potential terminological confusion associated with these Terms. First, their metaphoric origins contribute to a wide range of interpretations among the general public. Additionally, we observed the implicit duality of the concept of *greenhouse effect* which depending on the context refers to either 'enhanced (man-made) greenhouse effect' or 'natural greenhouse effect'. This duality is synonymous to that of *carbon dioxide* which incorporates both references 'polluting gas' and 'natural component of the atmosphere'.

Third, with accumulated knowledge about the use of *carbon* across text genres, we would want to explore the possibility of identifying and implementing automated solutions to detect deviant usages of scientific Terms in ordinary discourse. For example, we could use identified linguistic features of social network discourse to develop Natural Language Processing methodologies.

Finally, the changes in Twitter's internal policies that followed its acquisition by Elon Musk present an opportunity to conduct a comparative analysis of the discourse on this platform before and after the acquisition. Specifically, it would be relevant for environmental domain as the changes has affected the environmentally oriented audience, making many members stop their activity on the platform. As a consequence, we can expect to observe the effects of censorship such as the

emergence of new dominant opinion groups, change of the narrative, adoption of new lexicon, and development of new senses of environmental Terms.

Bibliography

Adelstein Andreína and Cabré M. Teresa (2002). The specificity of units with specialized meaning: polysemy as explanatory factor. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, 18, pp. 1–25.

Ahadh Abdhul, Binish Govind Vallabhasseri and Srinivasan Rajagopalan (2021). Text mining of accident reports using semi-supervised keyword extraction and topic modeling. *Process safety and environmental protection*, 155, pp. 455–465. issn: 0957-5820.

Allain Jean-François (2013). Durable. Vous avez dit durable ? *Traduire (Paris)*, 229, pp. 8–11. issn: 0395-773X.

Asiyanbi Adeniyi P. (2015). 'I don't get this climate stuff!' Making sense of climate change among the corporate middle class in Lagos. *Public Understanding of Science*, 24, pp. 1007–1024.

Bachram Heidi (2004). Climate fraud and carbon colonialism: the new trade in greenhouse gases. *Capitalism, nature, socialism*, 15:4, pp. 5–20. issn: 1045-5752.

Barr Stewart (2011). Climate forums: virtual discourses on climate change and the sustainable lifestyle. *Area*, 43:1, pp. 14–22. doi: https://doi.org/10.1111/j.1475-4762.2010.00958.x.

Bauer Laurie and Nation Paul (1993). Word Families. *International Journal of Lexicography*, 6:4, pp. 253–279.

BBC Science Focus (2020). *Climate change: should we change the terminology?* url: https://www.sciencefocus.com/news/climate-change-should-we-change-the-terminology/ (visited 21/03/2023).

Beacco Jean-Claude, Claudel Chantal, Doury Marianne, Petit Gèrard and Reboul-Touré Sandrine (2002). Science in media and social discourse: new

channels of communication, new linguistic forms. *Discourse studies*, 4:3, pp. 277–300. ISSN: 1461-4456.

BENESTAD Rasmus (2023). *The summary for policymakers of the Intergovernmental Panel on Climate Change sixth assessment reports synthesis*. URL: https://www.realclimate.org/index.php/archives/2023/04/the-summary-for-policymakers-of-the-intergovernmental-panel-on-climate-change-sixth-assessment-reports-synthesis/#more-24800 (visited 21/04/2023).

BERNERS-LEE Mike and CLARK Duncan (2010). *What is a carbon footprint?* URL: https://www.theguardian.com/environment/blog/2010/jun/04/carbon-footprint-definition (visited 02/07/2023).

BIESWANGER Markus (2013). 19. Micro-linguistic structural features of computer-mediated communication. In: *Pragmatics of Computer-Mediated Communication*. Ed. by Susan Herring, Dieter Stein and Tuija Virtanen. Berlin, Boston, De Gruyter Mouton, pp. 463–486. ISBN: 9783110214468. DOI: doi:10.1515/9783110214468.463.

BONADONNA Maria Francesca (2020). Using lexical functions to describe adjectives in terminography. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 26:1, pp. 7–32. ISSN: 0929-9971. DOI: https://doi.org/10.1075/term.00040.bon.

BOTTA Mariana (2013). La terminologie de l'environnement en vulgarisation scientifique : la famille lexicale de la régénération des forêts en portugais. *Equivalences*, 40:1, pp. 277–298. ISSN: 0751-9532.

BOUDET Hilary S. (2019). Public perceptions of and responses to new energy technologies. *Nature Energy*, 4:6, pp. 446–455. ISSN: 2058-7546.

BOWMAN Thomas E., MAIBACH Edward, MANN Michael E., MOSER Susanne C. and SOMERVILLE Richard C. J. (2009). Creating a Common Climate Language. *Science (American Association for the Advancement of Science)*, 324:5923, pp. 36–37. ISSN: 0036-8075.

BUDESCU David V, POR Han-Hui, BROOMELL Stephen B and SMITHSON Michael (2014). The interpretation of IPCC probabilistic statements around the world. *Nature climate change*, 4:6, pp. 508–512. ISSN: 1758-678X.

BUHR Katarina and HANSSON Anders (2011). Capturing the stories of corporations: A comparison of media debates on carbon capture and storage in

Norway and Sweden. *Global environmental change*, 21:2, pp. 336–345. ISSN: 0959-3780.

BUMPUS Adam G. (2011). The Matter of Carbon: Understanding the Materiality of tCO2e in Carbon Offsets. *Antipode*, 43:3, pp. 612–638. ISSN: 0066-4812.

BUSCH Timo and SHRIVASTAVA Paul (2011). *The global carbon crisis : emerging carbon constraints and strategic management options*. Greenleaf Publishing Sheffield, UK, 199 p. ISBN: 190609361 9781906093617.

CABEZAS-GARCÍA Melania and LEÓN-ARAÚZ Pilar (2023). Term and concept variation in climate change communication. *Translator (Manchester, England)*, 28:4, pp. 429–449. ISSN: 1355-6509. DOI: 10.1080/13556509.2023.2182168.

CABRÉ Maria Teresa and AZEVEDO Diego Napoleão Viana (2019). Teorias da terminologia: descrição, prescrição e explicação. *Cadernos de tradução*, 39:3. ISSN: 2175-7968.

CAMPBELL Benjamin, KHACHATRYAN Hayk, BEHE Bridget, DENNIS Jennifer and HALL Charles (2015). Consumer Perceptions of Eco-friendly and Sustainable Terms. *Agricultural and Resource Economics Review*, 44:1, pp. 21–34. DOI: 10.1017/S1068280500004603.

CAPSTICK Stuart, WHITMARSH Lorraine, POORTINGA Wouter, PIDGEON Nick and UPHAM Paul (2015). International trends in public perceptions of climate change over the past quarter century. *WIREs Climate Change*, 6:1, pp. 35–61. DOI: https://doi.org/10.1002/wcc.321.

CHANG Charlotte H., DESMUKH Nikhil R., ARMSWORTH Raul R. and MASUDA Yuta J. (2023). Environmental users abandoned Twitter after Musk takeover. *Trends in Ecology & Evolution*. DOI: https://doi.org/10.1016/j.tree.2023.07.002.

CHANG Kenneth (2008). *I hate organic*. URL: https://archive.nytimes.com/tierneylab.blogs.nytimes.com/2008/07/08/i-hate-organic/?mtrref=undefined&assetType=REGIWALL (visited 21/04/2023).

CHEN Chang-Feng, SHI Wen, YANG Jing and FU Hao-Huan (2021). Social bots' role in climate change discussion on Twitter: Measuring standpoints, topics, and interaction strategies. *Advances in climate change research*, 12:6, pp. 913–923. ISSN: 1674-9278.

Cody Emily M., Reagan Andrew J., Mitchell Lewis, Dodds Peter Sheridan and Danforth Christopher M. (2015). Climate Change Sentiment on Twitter: An Unsolicited Public Opinion Poll. *PloS one*, 10:8, e0136092–e0136092. issn: 1932-6203.

Collet Tanja (2021). What's a term? An attempt to define the term within the theoretical framework of text linguistics. *Linguistica Antverpiensia, New Series – Themes in Translation Studies*, 3.

Condamines Anne (1995). Terminology : new needs, new perspectives *. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 2:2, pp. 219–238. hal open archive: halshs-02131379.

Condamines Anne and Picton Aurélie (2014a). Des communiqués de presse du Cnes à la presse généraliste Vers un observatoire de la diffusion des termes. In P. Dury, J. C. d. Hoyos, J. Makri-Morel, F. Maniez, V. Renner and M. B. V. Diaz (eds.): *La néologie en langue de spécialité*. CRTT, pp. 165–188. hal open archive: halshs-01108940.

Condamines Anne and Picton Aurélie (2014b). Étude du fonctionnement des nominalisations déverbales dans un contexte de déspécialisation. *SHS web of conferences*, 8, pp. 697–711. issn: 2261-2424.

Cook John (2016). Countering climate science denial and communicating scientific consensus. In: *Oxford Research Encyclopedia*. Ed. by Hans Von Storch. United Kingdom, Oxford University Press. doi: 10.1093/acrefore/9780190228620.013.314.

Corner Adam, Shaw Chris and Clarke Jamie (2018). *Principles for effective communication and public engagement on climate change: A Handbook for IPCC authors*. Oxford: Climate Outreach. url: https://www.ipcc.ch/site/assets/uploads/2017/08/Climate-Outreach-IPCC-communications-handbook.pdf.

Costa Rute, Ramos Margarida, Salgado Ana, Carvalho Sara, Almeida Bruno and Silva Raquel (2022). Neoterm or neologism? A closer look at the determinologisation process. In: *Lexicography of Coronavirus-related Neologisms*. Ed. by Annette Klosa-Kückelhaus and Ilan Kernerman. Berlin, Boston, De Gruyter, pp. 237–260. isbn: 9783110798081. doi: doi:10.1515/9783110798081-012.

Cottey Alan (Aug. 2022). Climate and nature emergency: From scientists' warnings to sufficient action. *Public Understanding of Science*, 31:6, pp. 818–826. issn: 0963-6625. doi: 10.1177/09636625221100076.

Coulter Alice, Clegg Sue, Lyons Glenn, Chatterton Tim and Musselwhite Charles (2007). *Exploring public attitudes to personal carbon dioxide emission information*. Tech. rep. url: https://webarchive.nationalarchives.gov.uk/ukgwa/+/http:/www.dft.gov.uk/pgr/sustainable/climatechange/co2attitudes?page=1.

Coyle Kevin (2005). *Environmental literacy in America: What Ten Years of NEETF/Roper Research and Related Studies Say about Environmental Literacy in the U.S.* The National Environmental Education & Training Foundation, Washington, D.C.

Dabla-Norris Era, Helbling Thomas, Khalid Salma, Khan Hibah, Magistretti Giacomo, Sollaci Alexandre and Srinivasan Krishna (2023). *Public Perceptions of Climate Mitigation Policies: Evidence from Cross-Country Surveys*. International Monetary Fund. url: https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2023/02/07/Public-Perceptions-of-Climate-Mitigation-Policies-Evidence-from-Cross-Country-Surveys-528057.

Dancette Jeanne and L'Homme Marie-Claude (Jan. 2004). Building Specialized Dictionaries using Lexical Functions. *Linguistica Antverpiensia*, 3, pp. 113–132. doi: 10.52034/lanstts.v3i.107.

Davenport Coral (2018). *How Much Has 'Climate Change' Been Scrubbed From Federal Websites? A Lot.* url: https://www.nytimes.com/2018/01/10/climate/climate-change-trump.html (visited 21/04/2023).

Delavigne Valérie (2002). Le domaine aujourd'hui. Une notion à repenser. In: *Le traitement des marques de domaine en terminologie*. Danielle Candel. Paris, France. hal open archive: hal-00924228. url: https://hal.science/hal-00924228.

Delavigne Valérie (2003). Quand le terme entre en vulgarisation. In: *Terminologie et Intelligence artificielle*. Strasbourg, France, p. 80–91. url: https://hal.science/hal-00920636.

Dilling Lisa, Doney Scott C., Edmonds Jae, Gurney Kevin R., Harriss Robert, Schimel David, Stephens Britton and Stokes Gerald (2003). The role of carbon cycle observations and knowledge in carbon management. *Annual Review of Environment and Resources*, 28:1, pp. 521–558. doi: 10.1146/annurev.energy.28.011503.163443.

Dósa Katalin and Russ Rosemary S. (2020). Making sense of carbon footprints: how carbon literacy and quantitative literacy affects information gathering and decision-making. *Environmental Education Research*, 26:3, pp. 421–453. doi: 10.1080/13504622.2019.1569205.

Drouin Patrick, L'Homme Marie-Claude and Robichaud Benoıt (May 2018). Lexical Profiling of Environmental Corpora. In N. Calzolari, K. Choukri, C. Cieri, T. Declerck, S. Goggi, K. Hasida, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, S. Piperidis and T. Tokunaga (eds.): *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. Miyazaki, Japan, European Language Resources Association (ELRA). url: https://aclanthology.org/L18-1539.

Ha-Duong Minh, Nadaï Alain and Campos Ana Sofía (2009). A survey on the public perception of CCS in France. *International Journal of Greenhouse Gas Control*, 3:5, pp. 633–640. issn: 1750-5836. doi: https://doi.org/10.1016/j.ijggc.2009.05.003.

Dury Pascaline (2008). The rise of carbon neutral and compensation carbone: A diachronic investigation into the migration of vocabulary from the language of ecology to newspaper language and vice versa. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 14:2, pp. 230–248. issn: 0929-9971. doi: https://doi.org/10.1075/term.14.2.06dur.

EDF (2022). *Climate and Public Opinions International Observatory*. url: https://www.edf.fr/en/the-edf-group/climate-and-public-opinions-international-observatory/learnings (visited 06/04/2023).

EEA (2001). *Definition: carbon dioxide equivalent*. url https://www.eea.europa.eu/help/glossary/eea-glossary/carbon-dioxide-equivalent (visited 02/07/2023).

EEA (2023). *Climate*. url https://www.eea.europa.eu/en/topics/at-a-glance/climate (visited 05/04/2023).

Elgesem Dag, Steskal Lubos and Diakopoulos Nicholas (2015). Structure and Content of the Discourse on Climate Change in the Blogosphere: The Big Picture. *Environmental communication*, 9:2, pp. 169–188. issn: 1752-4032.

EPA (2022a). *Acid Rain Program*. url https://www.epa.gov/acidrain/acid-rain-program#overview (visited 15/04/2023).

EPA (2022b). *Organic Farming*. url https://www.epa.gov/agriculture/organic-farming#Background (visited 21/04/2023).

Ereaut Gill and Segnit Nat (2006). *Warm Words. How are we telling the climate story and can we tell it better?* London, the Institute for Public Policy Research.

European Commission (2022). *CO2 emissions of all world countries - JRC/IEA/PBL 2022 Report*. Ed. by M. Crippa, D. Guizzardi, M. Banja, E. Solazzo, M. Muntean, E. Schaaf, F. Pagani, F. Monforti-Ferrario, J. Olivier, R. Quadrelli, A. Risquez Martin, P. Taghavi-Moharamli, G. Grassi, S. Rossi, D. Jacome Felix Oom, A. Branco, J. San-Miguel-Ayanz and E. Vignati. Publications Office of the European Union, Luxembourg.

European Commission (2023a). *Carbon capture, storage and utilisation*. url: https://energy.ec.europa.eu/topics/oil-gas-and-coal/carbon-capture-storage-and-utilisation_en (visited 27/06/2023).

European Commission (2023b). *Consequences of climate change*. url: https://climate.ec.europa.eu/climate-change/consequences-climate-change_en (visited 04/05/2023).

Evans Ray (2009). *Thank God for carbon*. The Lavoisier Group. url: http://www.lavoisier.com.au/articles/climate-policy/science-and-policy/TGFC.pdf.

Faber Pamela and L'Homme Marie-Claude (eds.) (2014). *Lexical semantic approaches to terminology*. vol. 20. 2. John Benjamins Publishing Company, p. 161. doi: 10.1075/term.20.2.

Feinberg Matthew and Willer Robb (2011). Apocalypse Soon? Dire Messages Reduce Belief in Global Warming by Contradicting Just-World Beliefs. *Psychological Science*, 22:1, pp. 34–38. issn: 09567976, 14679280. url: http://www.jstor.org/stable/40984603.

Fillmore Charles J. (2006). Frame semantics. In: *Cognitive Linguistics: Basic Readings*. Ed. by Dirk Geeraerts. Berlin, New York, De Gruyter Mouton, pp. 373–400. isbn: 9783110199901. doi: 10.1515/9783110199901.373.

Fitzer E., Kochling K.-H., Boehm H. P. and Marsh H. (1995). Recommended terminology for the description of carbon as a solid (IUPAC Recommendations 1995). *Pure and Applied Chemistry*, 67:3, pp. 473–506. doi: 10.1351/pac199567030473.

Fletcher Jon and Downing Phil (2011). *Consumer understanding of green terms: a supplementary report on consumer responses to specific green terms – A report to the Department for Environment, Food and Rural Affairs*. Brook Lyndhurst & Icaro Consulting, for Defra, London.

Flottum Kjersti (2014a). A linguistic and discursive view on climate change discourse. *ASp*, 58, pp. 19–37. issn: 1246-8185.

Flottum Kjersti (2014b). Linguistic Mediation of Climate Change Discourse. *ASp*, 65, pp. 7–20. issn: 2108-6354.

Freixa Judit and Fernández-Silva Sabela (Aug. 2017). Chapter 7. Terminological variation and the unsaturability of concepts. In: *Multiple perspectives in terminological variation*. Ed. by Patrick Drouin, Aline Francœur, John Humbley and Aurélie Picton. Amsterdam/New York: John Benjamins, pp. 155–180. isbn: 978-9027223425. doi: 10.1075/tlrp.18.07fre.

Fuso Nerini Francesco, Fawcett Tina, Parag Yael and Ekins Paul (2021). Personal carbon allowances revisited. *Nature sustainability*, 4:12, pp. 1025–1031. issn: 2398-9629.

Gadema Zaina and Oglethorpe David (2011). The use and usefulness of carbon labelling food: A policy perspective from a survey of UK supermarket shoppers. *Food policy*, 36:6, pp. 815–822. issn: 0306-9192.

Gader Nabil, Lux-Pogodalla Veronika and Polguère Alain (Dec. 2012). Hand-Crafting a Lexical Network With a Knowledge-Based Graph Editor. In: *Proceedings of the 3rd Workshop on Cognitive Aspects of the Lexicon*. Mumbai, India, The COLING 2012 Organizing Committee, pp. 109–126. url: https://aclanthology.org/W12-5109.

Galisson Robert (1978). *Recherches de lexicologie descriptive, la banalisation lexicale: le vocabulaire du football dans la presse sportive, contribution aux*

recherches sur les langues techniques. Nathan université, information, formation. Paris, F. Nathan. URL: https://books.google.fr/books?id=E2I3AQAAIAAJ.

Gero Katy Ilonka, Liu Vivian, Huang Sarah, Lee Jennifer and Chilton Lydia B. (Oct. 2021). What Makes Tweetorials Tick: How Experts Communicate Complex Topics on Twitter. *Proc. ACM Hum.-Comput. Interact.*, 5:CSCW2. DOI: 10.1145/3479566.

Giusta Marina Della, Jaworska Sylvia and Greetham Danica Vukadinović (2021). Expert communication on Twitter: Comparing economists' and scientists' social networks, topics and communicative styles. *Public Understanding of Science*, 30:1, pp. 75–90. DOI: 10.1177/0963662520957252.

Glasgow Lindonne, Langaigne Benny, Thomas Carisha, Harvey Olando and Campbell Ezra Angella (2018). Public Knowledge and Attitudes towards Climate Change and Its Impacts on Ecosystems in Grenada. *American Journal of Climate Change*, 07, pp. 600–610.

Glavič Peter and Lukman Rebeka (2007). Review of sustainability terms and their definitions. *Journal of cleaner production*, 15:18, pp. 1875–1885. ISSN: 0959-6526.

Goldstein Benjamin, Gounaridis Dimitrios and Newell Joshua P. (2020). The carbon footprint of household energy use in the United States. *Proceedings of the National Academy of Sciences*, 117:32, pp. 19122–19130. DOI: 10.1073/pnas.1922205117.

Gorbunova Natal'ya N., Archakov Rustam A. and Akhtayeva Lalita A. (2020). Metaphorization as a basis for transterminologization in the sphere od subject terminological nominations. *Current Issues in Philology and Pedagogical Linguistics*, 3(2020), pp. 92–109. ISSN: 2079-6021.

Gorokhova Natalia (Feb. 2021). Determinologization And Transterminologization Processes In Modern Oil And Gas Discourse. In: *European Proceedings of Social and Behavioural Sciences*. European Publisher, pp. 329–335. DOI: 10.15405/epsbs.2021.02.02.42.

Gotkova Tomara and Chepurnykh Nikolay (2022). Public perception and usage of the term carbon: Linguistic analysis in an environmental social media corpus. *Psychology of language and communication*, 26:1, pp. 297–312. ISSN: 2083-8506.

Gotkova Tomara, Ingrosso Francesca, Mikhel Polina and Polguère Alain (2023). 'Defined_by' Relation as Structuring Principle of Terminologies. In: *Proceedings of the 2nd International Conference on Multilingual Digital Terminology Today (MDTT 2023)*. Lisbon, Portugal.

Gotkova Tomara and Shvets Alexander (2023). Key environmental lexicon extraction using generative transformer. In: *Proceedings of the 2nd International Conference on Multilingual Digital Terminology Today (MDTT 2023)*. Lisbon, Portugal.

Gowda M. V. Rajeev, Fox Jeffrey C. and Magelky Robin D. (1997). Students' Understanding of Climate Change: Insights for Scientists and Educators. *Bulletin of the American Meteorological Society*, 78:10, pp. 2232–2240. doi: https://doi.org/10.1175/1520-0477-78.10.2232.

Graff David, Kong Junbo, Chen Ke and Maeda Kazuaki (2003). English gigaword. *Linguistic Data Consortium, Philadelphia*, 4:1, p. 34.

Grinstein Amir, Koa Evan, Chen Stone, Sheldon Seth and Zik Ory (2018). Carbon innumeracy. *PloS one*, 13:5, e0196282–e0196282. issn: 1932-6203.

Guerris Manuel, Cuadros Jordi, González-Sabaté Lucinio and Serrano Vanessa (2020). Describing the public perception of chemistry on twitter. *Chemistry Education Research and Practice*, 21:3, pp. 989–999. issn: 1109-4028.

Halskov Jakob (2005). Probing the Properties of Determinologization: The DiaSketch. In: *Young Researchers at DCL*. Ed. by Steffen Leo Hansen. LAMBDA 29, pp. 39–63.

Hammond Geoffrey (Feb. 2007). Time to give due weight to the 'carbon footprint' issue. *Nature*, 445, p. 256. doi: 10.1038/445256b.

Hartikainen Hanna, Roininen Taneli, Katajajuuri Juha-Matti and Pulkkinen Hannele (2014). Finnish consumer perceptions of carbon footprints and carbon labelling of food products. *Journal of cleaner production*, 73, pp. 285–293. issn: 0959-6526.

Harvey Fiona (2023). *Eight things the world must do to avoid the worst of climate change*. url https://www.theguardian.com/environment/2023/mar/21/methane-to-food-waste-eight-ways-to-attempt-to-stay-within-15c (visited 20/04/2023).

Hassol Susan Joy (2008). Improving How Scientists Communicate About Climate Change. *Eos, Transactions American Geophysical Union*, 89:11, pp. 106–107. doi: https://doi.org/10.1029/2008EO110002.

Haunschild Robin, Leydesdorff Loet, Bornmann Lutz, Hellsten Iina and Marx Werner (2019). Does the public discuss other topics on climate change than researchers? A comparison of explorative networks based on author keywords and hashtags. *Journal of informetrics*, 13:2, pp. 695–707. issn: 1751-1577.

Hazen Robert M. (2019). *Symphony in C: Carbon and the Evolution of (Almost) Everything*. W. W. Norton & Company, 288 p. isbn: 0008292418 978-0008292416.

Hedin Björn, Grönborg Lucas and Johansson Gustav (Sept. 2022). Food Carbon Literacy: A Definition and Framework Exemplified by Designing and Evaluating a Digital Grocery List for Increasing Food Carbon Literacy and Changing Behavior. *Sustainability*, 14:19, pp. 1–16. url: https://ideas.repec.org/a/gam/jsusta/v14y2022i19p12442-d929702.html.

Helbig Christoph, Huether Jonas, Joachimsthaler Charlotte, Lehmann Christian, Raatz Simone, Thorenz Andrea, Faulstich Martin and Tuma Axel (2022). A terminology for downcycling. *Journal of Industrial Ecology*, 26:4, pp. 1164–1174. doi: https://doi.org/10.1111/jiec.13289.

Henry Adam Douglas (2000). Public Perceptions of Global Warming. *Human Ecology Review*, 7:1, pp. 25–30. issn: 10744827, 22040919. url: http://www.jstor.org/stable/24706893.

Hesselink Frits and van Kempen Peter Paul (May 1999). *Environmental education and training in Europe. Background paper for the European Enion Conference on EE&T in Europe*. Brussels, European Commission.

Hmielowski Jay D., Feldman Lauren, Myers Teresa A., Leiserowitz Anthony and Maibach Edward (2014). An attack on science? Media use, trust in scientists, and perceptions of global warming. *Public understanding of science*, 23:7. issn: 0963-6625.

Hollweg Karen, Tyler Jason, Bybee Rodger, Marcinkowski Tom, McBeth Bill and Zoido Paoblo (2011). *Developing a framework for assessing environmental literacy*. Washington, DC: North American Association for Environ-

mental Education. URL: https://cdn.naaee.org/sites/default/files/inline-files/devframewkassessenvlitonlineed.pdf.

Humbert-Droz Julie (2023). When Terms Become Neologisms: A Contribution to the Study of Neology from the Perspective of Determinologisation. *CLINA Revista Interdisciplinaria De Traducción Interpretación Y Comunicación Intercultural*, 9:7, pp. 135–158. URL: https://doi.org/10.14201/clina202391135158.

Humbert-Droz Julie and Picton Aurélie (2022). Du Higgs à la particule de Dieu : une analyse de la variation terminologique dans un contexte de déterminologisation. *SHS Web Conf.*, 138. DOI: 10.1051/shsconf/202213804005.

Humbley John (2009). Accounting for Term Formation. *Terminology Science & Research*, 20, pp. 5–20.

Humbley John (2018). 22. Term Formation And Neology. In: *An International Handbook*. Ed. by John Humbley, Gerhard Budin and Christer Laurén. Berlin, Boston, De Gruyter Mouton, pp. 437–452. ISBN: 9783110228014. DOI: 10.1515/9783110228014-022.

Ingrosso Francesca and Polguère Alain (Nov. 2015). How Terms Meet in Small-World Lexical Networks: The Case of Chemistry Terminology. In: *Terminology and Artificial Intelligence (TIA 2015)*. Thierry Poibeau and Pamela Faber. Granada, Spain, pp. 167–171. HAL open archive: halshs-01226185.

Iomdin Boris (2012). O 'nepravil'nom' ispol'zovanii terminov: mozhets li yazyk oshibat'sya? In: *Meanings, Texts, and Other Exciting Things*. Ed. by Juri D. Apresjan, Igor Boguslavsky, Marie-Claude L'Homme, Leonid Iomdin, Jasmina Milicevic, Alain Polguère and Leo Wanner. HAL CCSD, pp. 233–251.

IPCC (2005). *IPCC Special Report on Carbon Dioxide Capture and Storage*. Cambridge University Press. ISBN: ISBN-13 978-0-521-86643-9. URL: https://www.ipcc.ch/report/carbon-dioxide-capture-and-storage/.

IPCC (2007). *Fourth Assessment Report: Climate Change 2007: Working Group I Report: The Physical Science Basis*. Geneva: IPCC. URL: http://www.ipcc.ch/ipccreports/ar4-wg1.htm.

IPCC (2018). *Global warming of 1.5°C*. Ed. by Valerie Masson-Delmotte, P. Zhai, Hans-Otto Pörtner, Debra Roberts, Jim Skea, P. Shukla, Anna Pirani, Wilfran Moufouma-Okia, C. Péan, R. Pidcock, Sarah Connors, Robin Matthews,

Y. Chen, X. Zhou, Melissa Gomis, E. Lonnoy, T. Maycock, M. Tignor and MuhammadReza Tabatabaei. ISBN: 9781009157957. DOI: 10.1017/9781009157940.

IPCC (2021). *IPCC Sixth Assessment Report: Working Group 1: The Physical Science Basis*. Cambridge University Press. DOI: 10.1017/9781009157896.002. URL: https://www.ipcc.ch/report/ar6/wg1/.

IPCC (2022). *The evidence is clear: the time for action is now. We can halve emissions by 2030.* URL https://www.ipcc.ch/2022/04/04/ipcc-ar6-wgiii-pressrelease/ (visited 01/06/2023).

IPCC (2023a). *Annex I: Glossary, In: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Tech. rep. Geneva, Switzerland. DOI: doi:10.59327/IPCC/AR6-9789291691647.002.

IPCC (2023b). *Climate Change 2022 – Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. DOI: 10.1017/9781009325844.

Ipsos (2022). *Earth Day 2022: Public opinion on climate change*. Ipsos Global Advisor. URL: https://www.ipsos.com/sites/default/files/ct/news/documents/2022-04/Ipsos%5C%20-%5C%20Global%5C%20Advisor%5C%20-%5C%20Earth%5C%20Day%5C%202022%5C%20-%5C%20Release%5C%201.pdf.

ISO (2018). *ISO 14067:2018(en) Greenhouse gases — Carbon footprint of products — Requirements and guidelines for quantification*. URL: https://www.iso.org/obp/ui/en/#iso:std:iso:14067:ed-1:v1:en (visited 25/06/2023).

Ivanova Diana, Barrett John, Wiedenhofer Dominik, Macura Biljana, Callaghan Max and Creutzig Felix (2020). Quantifying the potential for climate change mitigation of consumption options. *Environmental research letters*, 15:9, pp. 93001–. ISSN: 1748-9326.

Ivanova Diana and Wood Richard (2020). The unequal distribution of household carbon footprints in Europe and its link to sustainability. *Global sustainability*, 3. ISSN: 2059-4798.

Jadot Yannick (2022). *Le programme des écologistes*. URL: https://www.jadot2022.fr/programme (visited 11/04/2023).

Jakucionyte-Skodiene Migle, Krikstolaitis Ricardas and Liobikiene Genovaite (2022). The contribution of changes in climate-friendly behaviour, climate change concern and personal responsibility to household greenhouse gas emissions: Heating/cooling and transport activities in the European Union. *Energy (Oxford)*, 246, pp. 123387–. issn: 0360-5442.

Jang S. Mo and Hart P. Sol (2015). Polarized frames on "climate change" and "global warming" across countries and states: Evidence from Twitter big data. *Global environmental change*, 32, pp. 11–17. issn: 0959-3780.

Jousse Anne-Laure and Bouveret Myriam (2003). Lexical functions to represent derivational relations in specialized dictionaries. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 9:1, pp. 71–98. issn: 0929-9971. doi: https://doi.org/10.1075/term.9.1.05jou.

Kageura Kyo (2002). *The Dynamics of Terminology: A descriptive theory of term formation and terminological growth*. John Benjamins. url: https://www.jbe-platform.com/content/books/9789027275363.

Khan Muhammad Qasim, Shahid Abdul, Uddin M. Irfan, Roman Muhammad, Alharbi Abdullah, Alosaimi Wael, Almalki Jameel and Alshahrani Saeed M. (2022). Impact analysis of keyword extraction using contextual word embedding. *PeerJ. Computer science*, 8, e967–e967. issn: 2376-5992.

Khan Osama A., Wasi Shaukat, Siddiqui Muhammad Shoaib and Karim Asim (2022). Keyword Extraction for Medium-Sized Documents Using Corpus-Based Contextual Semantic Smoothing. *Complexity (New York, N.Y.)*, 2022, pp. 1–8. issn: 1076-2787.

Khan Shair Ali (2016). The Distinction between Term and Word: A Translator and Interpreter Problem and the Role of Teaching Terminology. *Procedia - Social and Behavioral Sciences*, 232, pp. 696–704.

Kim Serena Y., Ganesan Koushik, Dickens Princess and Panda Soumya (2021). Public Sentiment toward Solar Energy—Opinion Mining of Twitter Using a Transformer-Based Language Model. *Sustainability (Basel, Switzerland)*, 13:5, pp. 2673–. issn: 2071-1050.

Kim Sojung Claire and Cooke Sandra L. (2018). Environmental framing on Twitter: Impact of Trump's Paris Agreement withdrawal on climate change and

ocean acidification dialogue. *Cogent environmental science*, 4:1, pp. 1532375–. ISSN: 2331-1843.

KITTREDGE Richard (1982). Sublanguages. *American Journal of Computational Linguistics*, 8:2, pp. 79–84. URL: https://aclanthology.org/J82-2006.

KLEINMAN Zoe (2023). *Twitter: Five ways Elon Musk has changed the platform for users*. URL: https://www.bbc.com/news/technology-64289251 (visited 20/07/2023).

KOLIC Blas, AGUIRRE-LÓPEZ Fabián, HERNÁNDEZ-WILLIAMS Sergio and GARDUÑO-HERNÁNDEZ Guillermo (2022). Quantifying the structure of controversial discussions with unsupervised methods: a look into the Twitter climate change conversation. *INET Oxford Working Paper No. 2022-23*. DOI: 10.48550/arXiv.2206.14501.

KOTEYKO Nelya (2010). Mining the internet for linguistic and social data: An analysis of 'carbon compounds' in Web feeds. *Discourse & society*, 21:6, pp. 655–674. ISSN: 0957-9265.

KOTEYKO Nelya (2012). Managing carbon emissions: A discursive presentation of 'market-driven sustainability' in the British media. *Language & communication*, 32:1, pp. 24–35. ISSN: 0271-5309.

KOTEYKO Nelya, THELWALL Mike and NERLICH Brigitte (2010). From Carbon Markets to Carbon Morality: Creative Compounds as Framing Devices in Online Discourses on Climate Change Mitigation. *Science Communication*, 32:1, pp. 25–54. DOI: 10.1177/1075547009340421.

KRIPPENDORFF K. (2004). *Content Analysis: An Introduction to Its Methodology*. Content Analysis: An Introduction to Its Methodology. Sage. ISBN: 9780761915454. URL: https://books.google.fr/books?id=q657o3M3C8cC.

KROSNICK Jon A., HOLBROOK Allyson, LOWE Laura and VISSER Penny S. (2006). The Origins and Consequences of democratic citizens' Policy Agendas: A Study of Popular Concern about Global Warming. *Climatic Change*, 77, pp. 7–43.

L'HOMME Marie-Claude (2015). Predicative Lexical Units in Terminology. In: *Recent Advances in Language Production, Cognition and the Lexicon*. Ed. by G. Bel-Enguix Nuria Gala Reinhard Rapp. Springer, pp. 75–93. ISBN: 978-3-319-08042-0. DOI: 10.1007/978-3-319-08043-7_6.

L'Homme Marie-Claude (Dec. 2003). Capturing the Lexical Structure in Special Subject Fields with Verbs and Verbal Derivatives. A Model for Specialized Lexicography. *International Journal of Lexicography*, 16:4, pp. 403–422. ISSN: 0950-3846. DOI: 10.1093/ijl/16.4.403.

L'Homme Marie-Claude (2016). Terminologie de l'environnement et Sémantique des cadres. *SHS Web of Conferences*, 27, pp. 05010–. ISSN: 2416-5182.

L'Homme Marie-Claude (2018). Le DiCoInfo – Méthodologie pour une nouvelle génération de dictionnaires spécialisés. *Traduire (Paris)*, pp. 78–103. ISSN: 0395-773X.

Láinez Ramos-Bossini Antonio Jesús and Tercedor Sánchez Maribel (2020). E-patients in Oncology: a corpus-based characterization of medical terminology in an online cancer forum. *Ibérica (Castellón de la Plana, Spain)*, 39:39, pp. 69–96. ISSN: 1139-7241.

Lakoff George (2010). Why it Matters How We Frame the Environment. *Environmental Communication*, 4:1, pp. 70–81. DOI: 10.1080/17524030903529749.

Lakoff George and Johnson Mark (2003). *Metaphors We Live By*. University of Chicago Press. ISBN: 9780226468013. URL: https://books.google.fr/books?id=zX8DPcwhENgC.

Law Jonathan and Rennie Richard (2020). *A Dictionary of Chemistry*. Oxford quick reference. Oxford, Oxford University Press. ISBN: 9780198841227. URL: https://books.google.fr/books?id=NeDQyAEACAAJ.

Ledouble Hélène (2020). Term circulation and conceptual instability in the mediation of science: Binary framing of the notions of biological versus chemical pesticides. *Discourse & Communication*, 14:5, pp. 466–488. DOI: 10.1177/1750481320917575.

Leiserowitz Anthony, Carman Jennifer, Buttermore Nicole, Neyens Liz, Rosenthal Seth, Marlon Jennifer, Schneider J.W. and Mulcahy Kelsey (2022). *International public opinion on climate change*. Report. New Haven. URL: https://climatecommunication.yale.edu/publications/international-public-opinion-on-climate-change-2022/.

Leiserowitz Anthony, Roser-Renouf Connie, Marlon Jennifer and Maibach Edward (2021). Global Warming's Six Americas: a review and recommenda-

tions for climate change communication. *Current opinion in behavioral sciences*, 42, pp. 97–103. ISSN: 2352-1546.

LEISEROWITZ Anthony, SMITH Nicholas and MARLON Jennifer (2010). *Americans' Knowledge of Climate Change*. Report. New Haven. URL: https://climatecommunication.yale.edu/publications/americans-knowledge-of-climate-change/.

LEÓN-ARAÚZ Pilar (Aug. 2017). Chapter 9. Term and concept variation in specialized knowledge dynamics. In: *Multiple perspectives in terminological variation*. Ed. by Patrick Drouin, Aline Francœur, John Humbley and Aurélie Picton. Amsterdam/New York: John Benjamins, pp. 213–258. ISBN: 978-9027223425. DOI: 10.1075/tlrp.18.07fre.

LEVY David L. and SPICER André (2013). Contested imaginaries and the cultural political economy of climate change. *Organization*, 20:5, pp. 659–678. DOI: 10.1177/1350508413489816.

LOEVE Sacha and BENSAUDE-VINCENT Bernadette (2017). The multiple signatures of carbon. In B. B. Vincent, S. Loeve, A. Nordmann and A. Schwarz (eds.): *Research objects in their technological setting*. History and philosophy of technoscience 10. Routledge. HAL open archive: halshs-01507185.

LOHMANN Larry (2001). Democracy or Carbocracy? Intellectual Corruption and the Future of the Climate Debate. In: *Briefing 24: Democracy or Carbocracy?* The Corner House.

LOPES Marta, ANTUNES Carlos Henggeler and JANDA Kathryn B. (2020). Chapter 1 - Energy and behaviour: Challenges of a low-carbon future. In M. Lopes, C. H. Antunes and K. B. Janda (eds.): *Energy and Behaviour*. Academic Press, pp. 1–15. ISBN: 978-0-12-818567-4. DOI: https://doi.org/10.1016/B978-0-12-818567-4.00030-2.

LORENZONI Irene, LEISEROWITZ Anthony, DORIA Miguel De Franca, POORTINGA Wouter and PIDGEON Nick F. (2006). Cross-National Comparisons of Image Associations with "Global Warming" and "Climate Change" Among Laypeople in the United States of America and Great Britain. *Journal of Risk Research*, 9:3, pp. 265–281. DOI: 10.1080/13669870600613658.

LOWE Thomas, BROWN Katrina, DESSAI Suraje, DE FRANÇA DORIA Miguel, HAYNES Kat and VINCENT Katharine (2006). Does tomorrow ever come?

Disaster narrative and public perceptions of climate change. *Public Understanding of Science*, 15:4, pp. 435–457. DOI: 10.1177/0963662506063796. HAL open archive: hal-00571094.

MACKEY Brendan, PRENTICE I. Colin, STEFFEN Will, HOUSE Joanna I., KEITH David Lindenmayer Heather and BERRY Sandra (2013). Untangling the confusion around land carbon science and climate change mitigation policy (vol 3, pg 552, 2013). *Nature climate change*, 3:9, pp. 847–847. ISSN: 1758-678X.

MANAHAN Stanley E. (2004). *Green Chemistry: Fundamentals of Sustainable Chemical Science and Technology*. ChemChar Research. ISBN: 9780974952208. URL: https://books.google.fr/books?id=NJGXAAAACAAJ.

MARÍN María José and REA Camino (2014). Researching legal terminology: a corpus-based proposal for the analysis of sub-technical legal terms. *ASp*, 66, pp. 61–82. ISSN: 1246-8185.

MARKEY Rob (2023). *From Carbon Confusion to Climate Clarity: The New World of Carbon Accounting and Reporting*. URL: https://www.netpromotersystem.com/insights/from-carbon-confusion-to-climate-clarity-the-new-world-of-carbon-accounting-and-reporting-nps-podcast/ (visited 02/07/2023).

MASIH Adven (2017). History and Philosophy of Environmental Science and Education and their interrelationship. *International Journal of Scientific Research*, 8:9. ISSN: 2229-5518.

MAYNARD Diana and BONTCHEVA Kalina (2015). Understanding climate change tweets: an open source toolkit for social media analysis. In: *EnviroInfo/ICT4S*. URL: https://api.semanticscholar.org/CorpusID:6870754.

MCCAFFREY Mark S. and BUHR Susan M. (2008). Clarifying Climate Confusion: Addressing Systemic Holes, Cognitive Gaps, and Misconceptions Through Climate Literacy. *Physical geography*, 29:6, pp. 512–528. ISSN: 0272-3646.

MCCARTHY Michael (2010). *Acid rain: An environmental crisis that disappeared off the radar*. URL: https://www.independent.co.uk/climate-change/news/acid-rain-an-environmental-crisis-that-disappeared-off-the-radar-2007838.html (visited 15/04/2023).

MCCLARAN Nikki, BEHE Bridget K., HUDDLESTON Patricia and FERNANDEZ R. Thomas (2020). Recycled or reclaimed? The effect of terminology on water

reuse perceptions. *Journal of Environmental Management*, 261, p. 110144. ISSN: 0301-4797. DOI: https://doi.org/10.1016/j.jenvman.2020.110144.

MEADOWS Donella H., MEADOWS Dennis L., RANDERS Jörgen and BEHRENS III William W. (1972). *The Limits to Growth. A Report for the Club of Rome's Project on the Predicament of Mankind*. New York, Universe Books. ISBN: 9780876631652. URL: http://catalog.hathitrust.org/api/volumes/oclc/2745197.html.

MEL'ČUK Igor (1995). *Introduction à la lexicologie explicative et combinatoire / Igor A. Mel'čuk, André Clas, Alain Polguère*. Universités francophones. Louvain-la-Neuve, Duculot. ISBN: 2-8011-1106-6.

MEL'ČUK Igor (2023). *General Phraseology. Theory and Practice*. Lingvisticæ Investigationes Supplementa 36. Amsterdam/Philadelphia, John Benjamins.

MEL'ČUK Igor and POLGUÈRE Alain (2018). Theory and Practice of Lexicographic Definition. *Journal of Cognitive Science*, 19:4, pp. 417–470. DOI: 10.17791/jcs.2018.19.4.417. URL: https://shs.hal.science/halshs-02089593.

MEL'ČUK Igor and POLGUÈRE Alain (July 2021). Les fonctions lexicales dernier cri. In S. Marengo (ed.): *La Théorie Sens-Texte. Concepts-clés et applications*. Dixit Grammatica. L'Harmattan, pp. 75–155. HAL open archive: hal-03311348. URL: https://hal.science/hal-03311348.

METAG Julia, FÜCHSLIN Tobias and SCHÄFER Mike S (2017). Global warming's five Germanys: A typology of Germans' views on climate change and patterns of media use and information. *Public Understanding of Science*, 26:4, pp. 434–451. ISSN: 0963-6625. DOI: 10.1177/0963662515592558.

METZ Bert, DAVIDSON Ogunlade, CONINCK Heleen de, LOOS Manuela and MEYER Leo (2005). IPCC Special Report on Carbon Dioxide Capture and Storage. *Cambridge University Press*. URL: http://www.ipcc.ch/pdf/special-reports/srccs/srccs_wholereport.pdf.

MEYER Ingrid and MACKINTOSH Kristen (2000). When terms move into our everyday lives: An overview of de-terminologization. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 6:1, pp. 111–138. ISSN: 0929-9971. DOI: https://doi.org/10.1075/term.6.1.07mey.

MEYER Ingrid, VARANTOLA Krista and MACKINTOSH Kristen (Aug. 1998). From Virtual Sex to Virtual Dictionaries: On the Analysis and Description of a

De-terminologized Word. In T. Fontenelle, P. Hiligsmann, A. Michiels, A. Moulin and S. Theissen (eds.): *Proceedings of the 8th EURALEX International Congress*. Liège, Belgium, Euralex, pp. 645–654. ISBN: 2-87233-091-7.

MIKHEL Polina (2022). *Multilingual study of chemistry lexicon at the interface of terminology and general language*. PhD thesis. Nancy, Université de Lorraine.

MILLER Tyler G. and SPOOLMAN Scott (2014). *Essentials of Ecology*. Cengage Learning. ISBN: 9781285982755. URL: https://books.google.fr/books?id=2sOiAgAAQBAJ.

MKONO Mucha and HUGHES Karen (2022). Sustainability paralysis in travel consumption: A Reddit study. *Tourism and Hospitality Research*. DOI: 10.1177/14673584221110370.

MOIRAND Sophie (2017). 2. Éléments de théorisation d'une linguistique du discours : L'exemple du discours sur les sciences dans les médias. *Modèles linguistiques*, 40, pp. 5–21. ISSN: 0249-6267.

MONROE Martha C., PLATE Richard R., OXARART Annie, BOWERS Alison and CHAVES Willandia A. (2019). Identifying effective climate change education strategies: a systematic review of the research. *Environmental Education Research*, 25:6, pp. 791–812. DOI: 10.1080/13504622.2017.1360842.

MOORE Darrel (2022). *Majority of smaller businesses don't understand 'carbon jargon'*. https://www.circularonline.co.uk/news/majority-of-smaller-businesses-dont-understand-carbon-jargon/ (visited 25/06/2023).

MORRISON Mark, DUNCAN Rod, SHERLEY Chris and PARTON Kathleen (2013). A comparison between attitudes to climate change in Australia and the United States. *Australasian Journal of Environmental Management*, 20:2, pp. 87–100. DOI: 10.1080/14486563.2012.762946.

MOSER Susanne C. (2010). Communicating climate change: history, challenges, process and future directions. *WIREs Climate Change*, 1:1, pp. 31–53. DOI: https://doi.org/10.1002/wcc.11.

MURRAY Aphra (2022). *Why Public Acceptance is Important for the Future of Carbon Capture and Utilization*. URL: https://earth.org/future-of-carbon-capture-and-utilization/ (visited 02/06/2023).

NASA (2015). *Keeping Up With Carbon*. URL: https://svs.gsfc.nasa.gov/10498 (visited 02/06/2023).

NASA (2016). *Understanding the Evolution of Life in the Universe.* URL: https://wmap.gsfc.nasa.gov/universe/uni_life.html (visited 02/07/2023).

NASA (2023a). *How do we know climate change is real?* URL: https://climate.nasa.gov/evidence/ (visited 05/04/2023).

NASA (2023b). *Scientific Consensus: Earth's Climate Is Warming.* URL: https://climate.nasa.gov/scientific-consensus/ (visited 27/04/2023).

Nerlich Brigitte (2012). 'Low carbon' metals, markets and metaphors: the creation of economic expectations about climate change mitigation. *Climatic change*, 110:1-2, pp. 31–51. ISSN: 0165-0009.

Nerlich Brigitte, Evans Vyvyan and Koteyko Nelya (2011). Low carbon diet: Reducing the complexities of climate change to human scale. *Language and cognition*, 3:1, pp. 45–82. ISSN: 1866-9808.

Nerlich Brigitte and Hellsten I. (2014). The Greenhouse Metaphor and the Footprint Metaphor. *Technikfolgenabschätzung – Theorie und Praxis*, 23:2. ISSN: 1619-7623.

Nerlich Brigitte and Jaspal Rusi (2013). UK media representations of Carbon Capture and Storage: Actors, frames and metaphors. *Metaphor and the Social World*, 3:1, pp. 35–53. ISSN: 2210-4070. DOI: https://doi.org/10.1075/msw.3.1.02ner.

Nerlich Brigitte and Koteyko Nelya (2009). Carbon Reduction Activism in the UK: Lexical Creativity and Lexical Framing in the Context of Climate Change. *Environmental communication*, 3:2, pp. 206–223. ISSN: 1752-4032.

Nerlich Brigitte and Koteyko Nelya (2010). Carbon Gold Rush and Carbon Cowboys: A New Chapter in Green Mythology? *Environmental communication*, 4:1, pp. 37–53. ISSN: 1752-4032.

Nerlich Brigitte, Koteyko Nelya and Brown Brian (2010). Theory and language of climate change communication. *WIREs Climate Change*, 1:1, pp. 97–110. DOI: https://doi.org/10.1002/wcc.2.

Newman Todd P. (2017). Tracking the release of IPCC AR5 on Twitter: Users, comments, and sources following the release of the Working Group I Summary for Policymakers. *Public Understanding of Science*, 26:7, pp. 815–825. DOI: 10.1177/0963662516628477.

NewScientist (2007). *Climate myths: Carbon dioxide isn't the most important greenhouse gas.* URL: https://www.newscientist.com/article/dn11652-climate-myths-carbon-dioxide-isnt-the-most-important-greenhouse-gas/ (visited 27/04/2023).

Nicolae Cristina and Delavigne Valérie (Sept. 2009). Naissance et circulation d'un terme : une histoire d'exoplanètes. In: *Journées de linguistique de corpus. Texte et corpus. Journées de linguistique de corpus 4.* Geoffrey Williams. Lorient, France, pp. 217–229. HAL open archive: hal-00924141.

Niranjan Ajit (2022). *Should I feel guilty about my carbon footprint?* URL: https://www.dw.com/en/should-i-feel-guilty-about-my-carbon-footprint/a-60080346 (visited 25/06/2023).

Nomoto Tadashi (2022). Keyword Extraction: A Modern Perspective. *SN computer science*, 4:1. ISSN: 2661-8907.

Nová Jana (July 2018). Terms Embraced by the General Public: How to Cope with Determinologization in the Dictionary? In J. Čibej, V. Gorjanc, I. Kosem and S. Krek (eds.): *Proceedings of the XVIII EURALEX International Congress: Lexicography in Global Contexts.* Ljubljana, Slovenia, Ljubljana University Press, Faculty of Arts, pp. 387–398. ISBN: 978-961-06-0097-8.

O'Connor Karen, Pimpalkhute Pranoti, Nikfarjam Azadeh, Ginn Rachel, Smith Karen L. and Gonzalez Graciela (2014). Pharmacovigilance on twitter? Mining tweets for adverse drug reactions. *AMIA. Annual Symposium proceedings*, 2014, pp. 924–933. ISSN: 1942-597X.

Olfe-Kräutlein Barbara, Armstrong Katy, Mutchek Michele, Cremonese Lorenzo and Sick Volker (2022). Why Terminology Matters for Successful Rollout of Carbon Dioxide Utilization Technologies. *Frontiers in climate*, 4. ISSN: 2624-9553.

Ollinger Sandrine, Polguère Alain, Chudy Yannick and Gaume Bruno (June 2020). Spiderlex et compagnie (Spiderlex & Co). In: *Actes de la 6e conférence conjointe Journées d'Études sur la Parole (JEP, 33e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition). Volume 4 : Démonstrations et résumés*

d'articles internationaux. Nancy, France, ATALA et AFCP, pp. 60–63. URL: https://aclanthology.org/2020.jeptalnrecital-demos.16.

OXFORD LANGUAGES (2023). *Word of the Year*. URL: https://languages.oup.com/word-of-the-year/ (visited 02/07/2023).

PANDEY Divya, AGRAWAL Madhoolika and PANDEY Jai Shanker (2011). Carbon footprint: current methods of estimation. *Environmental monitoring and assessment*, 178:1-4, pp. 135–160. ISSN: 0167-6369.

PATERSON Matthew and STRIPPLE Johannes (2012). Virtuous carbon. *Environmental politics*, 21:4, pp. 563–582. ISSN: 0964-4016.

PATT Anthony (2007). Assessing model-based and conflict-based uncertainty. *Global environmental change*, 17:1, pp. 37–46. ISSN: 0959-3780.

PEARCE Warren, HOLMBERG Kim, HELLSTEN Iina and NERLICH Brigitte (2014). Climate change on Twitter: topics, communities and conversations about the 2013 IPCC Working Group 1 report. *PloS one*, 9:4. ISSN: 1932-6203.

PEARCE Warren, NIEDERER Sabine, ÖZKULA S. Melisa and SÁNCHEZ QUERUBÍN Natalia (2019). The social media life of climate change: Platforms, publics, and future imaginaries. *Wiley interdisciplinary reviews. Climate change*, 10:2, e569–n/a. ISSN: 1757-7780.

PICTON Aurélie, CONDAMINES Anne and HUMBERT-DROZ Julie (Feb. 2021). Analyse diachronique du processus de déterminologisation. Une réflexion en diachronie courte en physique des particules. *Cahiers de Lexicologie*, 1:118, pp. 193–225. DOI: 10.48611/isbn.978-2-406-12006-3.p.0193. HAL open archive: halshs-03344283.

PIKE Cara, DOPPELT Bob and HERR Meredith (2010). *Climate Communications and Behavior Change: A Guide for Practitioners*. The Climate Leadership Initiative. URL: %5Curl%7Bhttps://scholarsbank.uoregon.edu/xmlui/bitstream/handle/1794/10708/ClimCommBehaviorChangeGuide.pdf?sequence=1%5C&isAllowed=y%7D.

POLGUÈRE Alain (Sept. 2014). From Writing Dictionaries to Weaving Lexical Networks. *International Journal of Lexicography*, 27:4, pp. 396–418. ISSN: 0950-3846. DOI: 10.1093/ijl/ecu017.

Polguère Alain (2016). *Lexicologie et sémantique lexicale. Notions fondamentales*. Paramètres. Les Presses de l'Université de Montréal. HAL open archive: hal-01413860.

Polguère Alain (2018). A Lexicographic Approach to the Study of Copolysemy Relations. *Russian Journal of Linguistics*, 22, pp. 788–820. DOI: 10.22363/2312-9182-2018-22-4-788-820.

Polguère Alain (2023). Les modificateurs redondants et leur rôle en terminologie : de l'amour physique aux masques de protection. In P. Frassi (ed.): *Phraséoologie et Terminologie*. Beihefte zur Zeitschrift für romanische Philologie. de Gruyter, pp. 83–104.

Praskievicz Sarah (2021). How the environment became global. *Anthropocene*, 35, p. 100305. ISSN: 2213-3054. DOI: https://doi.org/10.1016/j.ancene.2021.100305.

Proferes Nicholas, Jones Naiyan, Gilbert Sarah, Fiesler Casey and Zimmer Michael (2021). Studying Reddit: A Systematic Overview of Disciplines, Approaches, Methods, and Ethics. *Social media + society*, 7:2, pp. 205630512110190–. ISSN: 2056-3051.

Quirion Jean (2006). La localisation, palimpseste de l'aménagement terminologique ? Stratégies d'implantation terminologique et marketing. *Meta*, 51:4, pp. 824–837. DOI: https://doi.org/10.7202/014345ar.

Raffel Colin, Shazeer Noam, Roberts Adam, Lee Katherine, Narang Sharan, Matena Michael, Zhou Yanqi, Li Wei and Liu Peter J. (2020). Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. *Journal of Machine Learning Research*, 21:1. ISSN: 1532-4435.

Razeghi Manijeh (2019). The carbon atom. In: *The Mystery of Carbon*. 2053-2563. IOP Publishing, 1-1 to 1–12. ISBN: 978-0-7503-1182-3. DOI: 10.1088/2053-2563/ab35d1ch1.

Rey Alain (1979). *La terminologie : noms et notions / Alain Rey*. Que sais-je ? Paris, PUF. ISBN: 2-13-036047-5.

Rigouts Terryn Ayla, Hoste Véronique and Lefever Els (May 2018). A Gold Standard for Multilingual Automatic Term Extraction from Comparable Corpora: Term Structure and Translation Equivalents. In N. Calzolari, K. Choukri, C. Cieri, T. Declerck, S. Goggi, K. Hasida, H. Isahara, B. Maegaard,

J. Mariani, H. Mazo, A. Moreno, J. Odijk, S. Piperidis and T. Tokunaga (eds.): *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. Miyazaki, Japan, European Language Resources Association (ELRA). URL: https://aclanthology.org/L18-1284.

Rodhe Henning, Charlson Robert and Crawford Elisabeth (1997). Svante Arrhenius and the Greenhouse Effect. *Ambio*, 26:1, pp. 2–5. ISSN: 00447447, 16547209. URL: http://www.jstor.org/stable/4314542 (visited 08/17/2023).

Rondeau Guy (1984). *Introduction à la terminologie / Guy Rondeau*. 2e édition. Québec, G. Morin. ISBN: 2-89105-137-8.

Roxburgh Nicholas, Guan Dabo, Shin Kong Joo, Rand William, Managi Shunsuke, Lovelace Robin and Meng Jing (2019). Characterising climate change discourse on social media during extreme weather events. *Global environmental change*, 54, pp. 50–60. ISSN: 0959-3780.

Sadeghian Shadan, Wintersberger Philipp, Laschke Matthias and Hassenzahl Marc (2022). Designing Sustainable Mobility: Understanding Users' Behavior. In: *Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. AutomotiveUI '22. Seoul, Republic of Korea, Association for Computing Machinery, pp. 34–44. ISBN: 9781450394154. DOI: 10.1145/3543174.3546833.

Schaefer Mike S. (2012). Online communication on climate change and climate politics: a literature review. *Wiley interdisciplinary reviews. Climate change*, 3:6, pp. 527–543. ISSN: 1757-7780.

Schmid-Petri Hannah and Bürger Moritz (2022). The effect of misinformation and inoculation: Replication of an experiment on the effect of false experts in the context of climate change communication. *Public Understanding of Science*, 31:2, pp. 152–167. DOI: 10.1177/09636625211024550.

Schreier Margrit (2013). Qualitative Content Analysis. In: *The SAGE Handbook of Qualitative Data Analysis*. United Kingdom, SAGE Publications, Limited, pp. 170–183. ISBN: 1446208982.

Schuldt Jonathon P., Konrath Sara H. and Schwarz Norbert (2011). "Global Warming" or "Climate Change"? Whether the planet is warming depends on question wording. *The Public Opinion Quarterly*, 75:1, pp. 115–124. ISSN: 0033362X, 15375331. URL: http://www.jstor.org/stable/41288371.

Segerberg Alexandra and Bennett W. Lance (2011). Social Media and the Organization of Collective Action: Using Twitter to Explore the Ecologies of Two Climate Change Protests. *The communication review (Yverdon, Switzerland)*, 14:3, pp. 197–215. issn: 1071-4421.

Shah Miti, Seraj Sarah and Pennebaker James W. (2021). Climate Denial Fuels Climate Change Discussions More Than Local Climate-Related Disasters. *Frontiers in psychology*, 12, pp. 682057–682057. issn: 1664-1078.

Shi Wen, Fu Haohuan, Wang Peinan, Chen Changfeng and Xiong Jie (Feb. 2020). #Climatechange vs. #Globalwarming: Characterizing Two Competing Climate Discourses on Twitter with Semantic Network and Temporal Analyses. *IJERPH*, 17:3, pp. 1–22. url: https://ideas.repec.org/a/gam/jijerp/v17y2020i3p1062-d317907.html.

Shvets Alexander and Wanner Leo (2020). Concept Extraction Using Pointer–Generator Networks and Distant Supervision for Data Augmentation. In C. M. Keet and M. Dumontier (eds.): *Knowledge Engineering and Knowledge Management*. Cham, Springer International Publishing, pp. 120–135. isbn: 978-3-030-61244-3.

Simonchuk Vladislava and Romasheva Natalia (2021). Public perception of carbon capture (utilization) and storage projects: world experience and the situation in Russia. *E3S Web of Conferences*, 266, pp. 6008–. issn: 2267-1242.

Slovic Scott (June 2020). The Language of Warning: The World Scientists' Efforts to Communicate and the Challenge of Poignancy. *Ecocene: Cappadocia Journal of Environmental Humanities*, 1:1, pp. 44–51. doi: 10.46863/ecocene.33. url: https://ecocene.kapadokya.edu.tr/index.php/ecocene/article/view/5.

Statista (2023a). *Average monthly active users (MAU) of Reddit from April 2018 to January 2020*. url: https://www.statista.com/statistics/1324258/reddit-monthly-active-users/ (visited 21/07/2023).

Statista (2023b). *Number of Twitter users worldwide from 2019 to 2024*. url: https://www.statista.com/statistics/303681/twitter-users-worldwide/ (visited 20/07/2023).

Sterman John D. and Sweeney Linda B. (2002). Cloudy skies: assessing public understanding of global warming. *System Dynamics Review*, 18:2, pp. 207–240. doi: https://doi.org/10.1002/sdr.242.

Sterman John D. and Sweeney Linda B. (2007). Understanding public complacency about climate change: adults' mental models of climate change violate conservation of matter. *Climatic change*, 80:3-4, pp. 213–238. issn: 0165-0009.

Supran Geoffrey and Oreskes Naomi (2021). Rhetoric and frame analysis of ExxonMobil's climate change communications. *One Earth*, 4:5, pp. 696–719. issn: 2590-3322. doi: https://doi.org/10.1016/j.oneear.2021.04.014.

Talmy Leonard (2000). *Toward a cognitive semantics . Volume 1, Concept structuring systems / Leonard Talmy*. Language, Speech, and Communication. Cambridge, Mass. London, The MIT Press. isbn: 0-262-20120-8.

Temmerman Rita (2000). *Towards New Ways of Terminology Description: The sociocognitive approach*. John Benjamins. url: https://www.jbe-platform.com/content/books/9789027298638.

The European Investment Bank (2020). *2019-2020 EIB climate survey*. url: https://www.eib.org/en/surveys/2nd-climate-survey/new-years-resolutions.htm (visited 04/26/2023).

The Guardian (2021). *Guardian and Observer style guide*. url: https://www.theguardian.com/guardian-observer-style-guide-c (visited 19/04/2023).

The Guardian (2022). *World is on 'highway to climate hell', UN chief warns at Cop27 summit*. url: https://www.theguardian.com/environment/2022/nov/07/cop27-climate-summit-un-secretary-general-antonio-guterres (visited 24/04/2023).

Thornton Alex (2009). *Public attitudes and behaviours towards the environment - tracker survey: A report to the Department for Environment, Food and Rural Affairs*. TNS. Defra, London.

Toupin Rémi, Millerand Florence and Larivière Vincent (2022). Who tweets climate change papers? Investigating publics of research through users' descriptions. *PloS one*, 17:6, e0268999–e0268999. issn: 1932-6203.

Treen Kathie, Williams Hywel, O'Neill Saffron and Coan Travis G. (2022). Discussion of Climate Change on Reddit: Polarized Discourse or Deliberative Debate? *Environmental communication*, 16:5, pp. 680–698. issn: 1752-4032.

Trifonov D.N., Trifonov V.D., Glebov O.A. and Poluyan I.V. (1982). *Chemical Elements: How They Were Discovered*. Mir Publishers. isbn: 9785030007786. url: https://books.google.fr/books?id=4WnkugEACAAJ.

Trouwloon Danick, Streck Charlotte, Chagas Thiago and Martinus Glenpherd (2023). Understanding the Use of Carbon Credits by Companies: A Review of the Defining Elements of Corporate Climate Claims. *Global challenges*, 7:4. issn: 2056-6646.

Trumbo Craig W. and Shanahan James (2000). Social research on climate change: where we have been, where we are, and where we might go. *Public understanding of science (Bristol, England)*, 9:3, pp. 199–204. issn: 0963-6625.

Twitter Developer Platform (2023). *Developer Agreement and Policy*. url: https://developer.twitter.com/en/developer-terms/agreement-and-policy (visited 21/07/2023).

Twyman Chasca, Smith Thomas Aneurin and Arnall Alex (2015). What is carbon? Conceptualising carbon and capabilities in the context of community sequestration projects in the global South. *WIREs Climate Change*, 6:6, pp. 627–641. doi: https://doi.org/10.1002/wcc.367.

U.S Climate Change Science Program (2009). *Climate Literacy: The Essential Principles of Climate Science*. Climate Change Science Program (U.S.) url: https://www.climate.gov/teaching/climate.

UN (2023). *What Is Climate Change?* url: https://www.un.org/en/climatechange/what-is-climate-change (visited 27/04/2023).

UNEP (2021). *Consequences of climate change*. https://www.unep.org/explore-topics/climate-action/what-we-do/climate-action-note/state-of-climate.html (visited 05/04/2023).

Ungar Sheldon (2000). Knowledge, ignorance and the popular culture: climate change versus the ozone hole. *Public Understanding of Science*, 9:3, pp. 297–312. doi: 10.1088/0963-6625/9/3/306.

Union of Concerned Scientists (1992). *1992 World Scientists' Warning to Humanity*. url: https://www.ucsusa.org/resources/1992-world-scientists-warning-humanity#:~:text=We%20the%20undersigned%2C%20senior%20members,not%20to%20be%20irretrievably%20mutilated. (visited 21/03/2023).

Vargas Elodie (2009). Discours de vulgarisation à travers différents médias ou les tribulations des termes scientifiques. *Les Cahiers de l'ILCEA*, 11. issn: 1639-6073. doi: https://doi.org/10.4000/ilcea.217.

Veltri Giuseppe Alessandro and Atanasova Dimitrinka (2017). Climate change on Twitter: Content, media ecology and information sharing behaviour. *Public Understanding of Science*, 26, pp. 721–737. URL: https://api.semanticscholar.org/CorpusID:663113.

Wallquist Lasse, Visschers Vivianne H.M. and Siegrist Michael (2009). Lay concepts on CCS deployment in Switzerland based on qualitative interviews. *International journal of greenhouse gas control*, 3:5, pp. 652–657. ISSN: 1750-5836.

Walter Stefanie, Lörcher Ines and Brüggemann Michael (Apr. 2019). Scientific networks on Twitter: Analyzing scientists' interactions in the climate change debate. *Public Understanding of Science*, 28:6, pp. 696–712. DOI: 10.1177/0963662519844131.

Washington Post (2023). *Scientists have found a perfect illustration of how the climate is spiraling 'out of control'*. URL: https://www.washingtonpost.com/news/energy-environment/wp/2016/07/28/these-climate-spirals-perfectly-illustrate-the-human-hand-in-global-climate-change/ (visited 04/04/2023).

Watkins James and Durning Bridget (2012). Carbon definitions and typologies in environmental impact assessment: greenhouse gas confusion? *Impact assessment and project appraisal*, 30:4, pp. 296–301. ISSN: 1461-5517.

Weber Elke U and Stern Paul C (2011). Public Understanding of Climate Change in the United States. *The American psychologist*. Psychology and Global Climate Change 66:4, pp. 315–328. ISSN: 0003-066X.

Whitmarsh Lorraine (2009). What's in a name? Commonalities and differences in public understanding of "climate change" and "global warming". *Public understanding of science (Bristol, England)*, 18:4, pp. 401–420. ISSN: 0963-6625.

Whitmarsh Lorraine, Seyfang Gill and O'Neill Saffron (2011). Public engagement with carbon and climate change: To what extent is the public 'carbon capable'? *Global Environmental Change*, 21:1, pp. 56–65. ISSN: 0959-3780. DOI: https://doi.org/10.1016/j.gloenvcha.2010.07.011.

Wibeck Victoria, Hansson Anders, Anshelm Jonas, Asayama Shinichiro, Dilling Lisa, Feetham Pamela M., Hauser Rachel, Ishii Atsushi and Sugiyama Masahiro (Nov. 2017). Making sense of climate engineering: a focus group

study of lay publics in four countries. *Climatic Change*, 145:1, pp. 1–14. DOI: 10.1007/s10584-017-2067-0.

WIEDMANN T. and MINX J. (2007). A definition of 'carbon footprint'. *CC Pertsova, Ecological Economics Research Trends*, 2, pp. 55–65. URL: http://scholar.google.co.uk/scholar.bib?q=info:h3OopwfI5cMJ:scholar.google.com/&output=citation&hl=en&as_sdt=0,5&ct=citation&cd=0.

WILLIAMS Hywel T.P., MCMURRAY James R., KURZ Tim and HUGO LAMBERT F. (2015). Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Global environmental change*, 32, pp. 126–138. ISSN: 0959-3780.

WOOD Dan and VEDLITZ Arnold (2007). Issue Definition, Information Processing, and the Politics of Global Warming. *American journal of political science*, 51:3, pp. 552–568. ISSN: 0092-5853.

WU Pei-Hao, YU Avon, TSAI Ching-Wei, KOH Jia-Ling, KUO Chin-Chi and CHEN Arbee L. P. (2020). Keyword extraction and structuralization of medical reports. *Health information science and systems*, 8:1, pp. 18–18. ISSN: 2047-2501.

WÜSTER Eugen (1991). *Einführung in die Allgemeine Terminologielehre und terminologische Lexikographie / Eugen Wüster ; mit einem Vorwort von Richard Baum*. 3. Aufl. Abhandlungen zur Sprache und Literatur. Bonn, Romanistischer Verlag. ISBN: 3-924888-48-5.

WWF (2014). *Busting the myths. Debunking myths about renewable energy*. URL: https://wwfint.awsassets.panda.org/downloads/myth_buster_brochure.pdf (visited 27/04/2023).

WWF (2023). *What are climate change and global warming?* URL: https://www.wwf.org.uk/climate-change-and-global-warming (visited 21/03/2023).

XU Wentao and SASAHARA Kazutoshi (2022). Characterizing the roles of bots on Twitter during the COVID-19 infodemic. *Journal of computational social science*, 5:1, pp. 591–609. ISSN: 2432-2717.

YAO Hong, GAO Xinyu, GUO Jingbo, WANG Hui, ZHANG Liang, FAN Liru, JIA Fangxu, GUO Jianhua and PENG Yongzhen (2022). Contribution of nitrous oxide to the carbon footprint of full-scale wastewater treatment plants and mitigation strategies- a critical review. *Environmental pollution (1987)*, 314, pp. 120295–120295. ISSN: 0269-7491.

Yu Pei-Shan, Hoepner Andreas and Adamsson Hampus (Jan. 2016). Towards a Carbon Data Science. *SSRN Electronic Journal*. doi: 10.2139/ssrn.2715101.

Zarrabeitia-Bilbao Enara, Rio-Belver Rosa-María, Alvarez-Meaza Izaskun and Martínez de Alegría-Mancisidor Itziar (2022). World Environment Day: Understanding Environmental Programs Impact on Society Using Twitter Data Mining. *Social indicators research*, 164:1, pp. 263–284. issn: 0303-8300.

Zimmer Ben (2007). *Compounding Carbon Confusion*. url: https://blog.oup.com/2007/07/carbon/ (visited 20/08/2023).

Zinke Laura (2020). The colours of carbon. *Nature reviews. Earth & environment*, 1:3, pp. 141–141. issn: 2662-138X.

Appendix A

List of keywords

Keywords marked with an asterisk are non-autonomous keywords (see 4.1.3), which means that they are not self-sufficient entities and require additional complementary keywords to retrieve relevant environmental texts.

acidification	aquatic ecosystem
*aerosol	arctic warming
aerosol emission	*atmosphere
afforestation	atmospheric carbon
*air	atmospheric carbon dioxide
air pollutant	atmospheric methane
air pollution	atmospheric pollution
*air temperature	atmospheric temperature
*anthropogenic	atmospheric warming
anthropogenic aerosol	*benign
anthropogenic carbon	benign chemical
anthropogenic carbon dioxide	bio-based
anthropogenic chemical	bioaccumulative chemical
anthropogenic climate	biodiversity
anthropogenic emission	biodiversity conservation
anthropogenic global warming	biodiversity hotspot
anthropogenic greenhouse gas	biodiversity loss
anthropogenic warming	bioeconomy

bioenergy
biomass
biomethane
biosphere
black carbon
*carbon
carbon balance
carbon budget
carbon capture
carbon concentration
carbon credit
carbon cycle
*carbon dioxide
carbon dioxide capture
carbon dioxide emission
carbon dioxide equivalent
carbon dioxide pollution
carbon dioxide removal
carbon emission
carbon energy
carbon exchange
carbon flux
carbon footprint
carbon management
*carbon monoxide
carbon pollution
carbon sequestration
carbon sink
carbon source
carbon tax
carbon-emitting
carbon-free
carbon-intensive
carbon-neutral
*ch4
*chemical
chemical pollution
circular economy
*clean
clean energy
*climate
climate action
climate adaptation
climate change
climate crisis
climate emission
climate event
climate impact
climate mitigation
climate model
climate policy
climate pollutant
climate science
climate system

climate warming

climate zone

*co2

co2e

*coal

coal consumption

coastal ecosystem

*conservation

contamination

cryosphere

dangerous chemical

decarbonisation

decarbonization

deforestation

*degradation

destruction of ozone

downcycling

*drought

e-waste

*earth

ecological degradation

ecological diversity

ecological drought

ecology

ecosystem

ecosystem degradation

*emission

emission reduction

*emit

endangered species

*energy

energy use

the environment

environmental challenge

environmental change

environmental degradation

environmental policy

environmental pollution

environmental sustainability

environmentally friendly

environmentally sustainable

*erosion

extreme event

extreme temperature

extreme weather

extreme weather event

*forest

forest loss

fossil fuel

fossil fuel combustion

fossil fuel emission

fracking

ghg

ghg emission

glacial ice
global average temperature
global biodiversity
global climate
global climate change
global climate model
global emission
global ocean
global sea level
global sea level rise
global temperature
global warming
*green
green chemistry
green economy
green energy
green policy
greenhouse effect
grenhouse gas
greenhouse gas emission
*habitat
habitat degradation
habitat loss
harmful chemical
hazardous chemical
hazardous waste
*heat
heat wave
heat-trapping gas
*ice
ice cover
ice sheet
industrial chemical
inorganic carbon
*land
land cover
land degradation
land pollution
land surface temperature
land-use
*landfill
landfill gas
landfill waste
*litter
loss of biodiversity
low-carbon
low-carbon economy
*marine
marine ecosystem
marine environment
*marine life
marine litter
marine pollution
marine species

*meteorological
*methane
methane emission
microplastic
*mitigation
mobile chemical
nanoplastic
natural ecosystem
*nature
*non-toxic
non-toxic chemical
non-toxic environment
*o3
*ocean
ocean acidification
ocean temperature
*organic
organic carbon
*ozone
ozone concentration
ozone depletion
ozone emissions
ozone hole
ozone layer
ozone level
ozone pollution
ozone-depleting substance
particulate matter
*peatland
persistent chemical
*plastic
plastic litter
plastic pollution
plastic waste
*pm
pollutant emission
pollution
*precipitation
precipitation rate
*radiation
*rainforest
recycling
reduce emissions
reforestation
regional climate
renewable
renewable electricity
renewable energy
rising temperature
*safe
safe chemical
sea level
sea surface temperature
*smog

*so2
soil degradation
soil pollution
solar radiation
solid waste
*species
species diversity
species extinction
surface temperature
surface weather
sustainability
sustainable chemical
sustainable consumption
sustainable development
sustainable economy
synthetic chemical
*temperature
temperature rise
threatened species
total emissions
*toxic
toxic chemical
toxic waste
upcycling
*vegetation
*warming
warming climate
*waste
waste landfill
waste management
*wastewater
water pollution
water vapour
*weather
*weathering
wildfire
zero-carbon
zero-emission

Index of notions