



HAL
open science

Climate change and Twitter in Belgium

Andrea Pizarro Pedraza, Louise-Amélie Cougnon, Andrea Catellani, Yuliya Samofalova, Olivier de Schutter

► **To cite this version:**

Andrea Pizarro Pedraza, Louise-Amélie Cougnon, Andrea Catellani, Yuliya Samofalova, Olivier de Schutter. Climate change and Twitter in Belgium. Université Catholique de Louvain. 2022. <hal-05035637>

HAL Id: hal-05035637

<https://hal.science/hal-05035637v1>

Submitted on 15 Apr 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization



Climate change and Twitter in Belgium

Report for the project *202CM Overcoming Obstacles and Disincentives to Climate Change Mitigation: A cross-cutting approach by human and social sciences* (funded by Solstice, JPI Climate)

©Consortium 202CM January 2022

Citation: Pizarro Pedraza, A., Coughon, L.-A., Catellani, A., Samofalova, Y. & De Schutter, O. (2022). *Climate change and Twitter in Belgium*. Report for the 202CM project, SocArXiv papers.

CONTENTS

1. Introduction.....	2
1.1. Framework and research goals.....	2
1.2. Climate change in social media.....	2
1.3. The linguistic expression of obstacles.....	3
2. Methodology.....	4
2.1. Corpus collection.....	4
2.2. Corpus description.....	5
2.2. Analytical methods: corpus linguistics.....	1
3. Results.....	2
3.1. General trends in the corpus.....	2
3.1.1. Most typical terms of the corpus.....	2
3.1.2. Most typical multiword expressions of the corpus.....	4
3.2. Trends related to food, transport and energetic efficiency.....	5
3.2.1. Definition and scope of the three axes.....	5
3.2.2. General distribution.....	6
3.2.1. Trends for Food terms.....	9
3.2.2. Trends for Transport terms.....	12
3.2.3. Trends for Energetic efficiency terms.....	15
3.3. The expression of obstacles in taking action against climate change.....	18
4. Summary: key results of the report.....	23
5. References.....	25

1. INTRODUCTION

1.1. FRAMEWORK AND RESEARCH GOALS

This report is part of the project *202CM Overcoming Obstacles and Disincentives to Climate Change Mitigation: A cross-cutting approach by human and social sciences*, funded by Solstice (JPI Climate). The project “has the theoretical and transformative aim to improve the scientific understanding of obstacles and disincentives to climate change mitigation and to understand how multimodal devices can operate on them, through the interdisciplinary collaboration between linguistics and social psychological approaches”. This report belongs specifically to the WP2, which focuses particularly on the analysis of text and image from social media, such as Twitter, Reddit and Instagram. Here, we present the results of the linguistic analysis of Twitter data in French-speaking Belgium related to climate change.

The goals of this analysis are to understand “how climate change is depicted and communicated in Twitter; what are the obstacles to the adoption of attitudes and behaviours of climate change mitigation: forms of resistance, deformation, and denial; and what textual sources are mostly read, cited and commented on by laypeople and authorities online.” We will tackle these questions from a linguistic perspective, and particularly with a combination of methods from corpus linguistics and discourse analysis (which will be explained in detail in section 2.). This will translate very specifically in the following objectives:

Objective 1: Collect a corpus of Belgian French, Climate change-related tweets

Objective 2: Analyze the distribution of the data across segments, time and particular accounts (most active, most mentioned)

Objective 3: Analyze the presence and distribution of several important axes related to Climate change that have a crucial impact on individual life (food, transport, energetic efficiency)

Objective 4: Extract and analyze the expressions of difficulty in the corpus at different linguistic levels

In order to better understand these goals, part 1.2 will summarize previous findings related to climate change in social media. Then part 2.2 will deal with the concept of obstacle and how it can be tackled from a linguistic point of view.

1.2. CLIMATE CHANGE IN SOCIAL MEDIA

Social media have become extremely popular as a source of data for linguistic studies, which is no surprise since they provide an immense amount of written texts that can be accessed in a relatively easy way. People’s textual productions in social media are unprompted, which is

particularly valuable, since the expressed views, and the ways in which they are expressed, exist independently of the researchers and their research goals.

The large amounts of data available online have also facilitated linguistic studies about the climate change debate. Salway (2017) states that “[r]ecent collections of English-language texts alone suggest 30,000 newspaper articles, 2.8million tweets, and 150,000 blog posts per month broadly related to climate change (Olteanu et al. 2015; Salway et al. 2016)”. These amounts of data were not imaginable before. Logically, new analytical methods have also been adopted, namely data-driven approaches. Salway (2017:156) summarizes the techniques in two main groups, those that focus on annotation of large groups of texts (known as corpora) for classification purposes, for instance, and those that focus on finding and analyzing linguistic patterns. Typically, the goal of the first is to automatically detect climate change-related texts from other topics, or classify different climate change subtopics. This has been done by using ad hoc lists of hashtags (Kirilenko and Stepchenkova 2014, Pearce et al 2014) or terms (Mayer 2012, Jang and Hart 2015), but also through unsupervised techniques such as topic modeling (Tvinnereim and Fløttum 2015). For the second group of studies, the goal is to find patterns through corpus-based discourse analysis, which calculates frequency of appearance of individual words (such as keywords) and clusters of words (such as collocations) (as in Koteyko 2010, Grundmann and Krishnamurthy 2010, Wild et al. 2013, Grundmann and Scott 2014).

In this report, the focus is particularly on Twitter, a social media that has been used in some of the studies mentioned before. We have used a combination of the mentioned techniques: ad hoc lists of terms and hashtags for classification, and then corpus linguistic techniques for analysis. The details will be explained in 2.

1.3. THE LINGUISTIC EXPRESSION OF OBSTACLES

Besides the collection of a Twitter corpus of climate change, an important goal of this project is to better “understand the obstacles and disincentives to climate change mitigation” in this particular corpus. Specifically, the goal is to understand why individuals find it difficult to take concrete actions against climate change, and what very concrete obstacles they encounter.

It is known that Climate Change is generally an accepted reality for institutions, companies and individuals in Belgium, however implementing actions against climate change is difficult in daily life (Bouman et al., 2021; Bourg et al., 2006).

Although difficulty is not a linguistic term, it is linked to well-established concepts such as refusal, which is a speech act that consists of refusing a request. When refusing to do something, speakers usually also give an excuse to justify their reaction. This is because refusal is considered a dispreferred speech act, which leaves the participants in an awkward, face-threatening position. Obstacles are often expressed as reasons for refusals (Ifert Johnson, Roloff & Riffée 2004).

In our case, public discourses about climate change include directive messages for individuals to take actions, that can be interpreted as “societal requests”, as in this slogan by Ecoconso *Arrête d'en faire des tonnes*, which demands the public to reduce their CO2 impact. It is the refusal of this “public” or societal request (acting against climate change) that leads to the expression of the difficulties it entails.

2. METHODOLOGY

2.1. CORPUS COLLECTION

For this study, we used the Twitter collection methodology developed by Cougnon and De Viron (2021). Their method is designed to collect tweets from a particular country, which is often challenging in social media, since users rarely display their geographical location. For the case of French-speaking Belgium, Cougnon and De Viron had selected Twitter accounts from Belgian political figures and media channels, and used them to collect accounts of their followers. The reasoning behind this method is that the content produced by these Belgian influencers is mostly relevant for Belgians. Belgian accounts of lay people were then used to collect their tweets. The corpus is regularly updated and therefore always growing. The corpus is divided in three social segments (politicians, media, population) and has other metadata such as date, attached links, mentions, or retweet status.

In order to reduce the general corpus to a subcorpus relevant for our goals, we needed to restrict it by topic, since our goal is to analyze how users communicate about climate change. For that purpose, we proceeded to create a list of expressions that are unequivocally related to climate change, so that we could use them for topical restriction in the creation of the subcorpus. We did this through an internal API, in which several parameters of the query can be modified, namely the expressions that need to be present in the retrieved tweets. We started with a basic list of climate change-related terms. It is important to note that we kept the topical selection broad, in the sense that the terms are related to climate change in general (causes, consequences, effects... of climate change), but also to ecology. The resulting corpus was then checked for relevance and we discarded irrelevant terms. A term was considered irrelevant if it was present in tweets that were not strictly related to climate change.

We also applied a keyword analysis to this first set of tweets, and to the list of present hashtags. This type of analysis determines the expressions that are significantly more typical of a particular corpus, in comparison to a general corpus, and can be useful to find terms that are related to a topic, like climate change. We then used the resulting keyword and hashtag list to retrieve a new corpus through the API. The same procedure was repeated again: the corpus was checked for relevance, so that irrelevant query terms could be discarded, and keyword and key-hashtag analyses were applied. At this stage we considered that the retrieved tweets were vastly related to climate change and we proceeded to collect the final corpus.

2.2. CORPUS DESCRIPTION

The Belgian French Corpus of Climate Change (BFCCC) (as used in our analyses) was collected in September 2021. It is composed of 385 977 tweets. The distribution among the segments and through time is unbalanced: on the one hand, most of the tweets are from the population accounts¹ (Fig. 1), and on the other, most tweets were posted from 2016 onwards (Fig. 2).

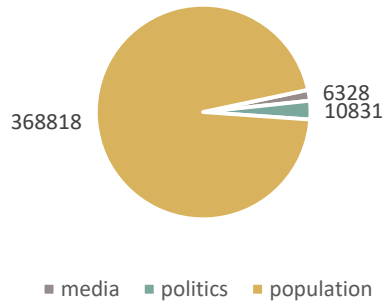


Figure 1 Distribution of tweets across the three different social segments in the corpus (media, politics, population, in raw numbers)

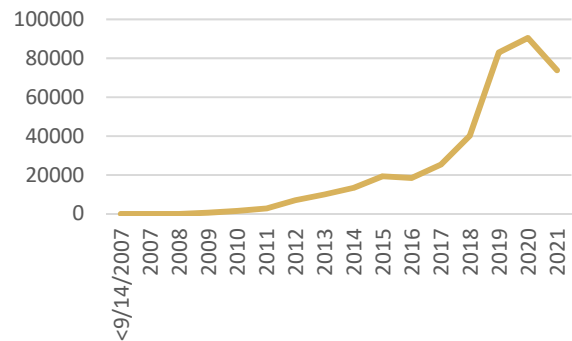


Figure 2 Evolution of tweets posted from 2007 until collection date (September 2021, raw numbers)

It is important to mention that the scarcity of tweets in the early years is due to a methodological limitation in the number of tweets that can be collected back in time, and cannot be taken as an indication of the actual amount of tweets produced about this topic.

Another interesting aspect to describe is the number of tweets per account (Fig. 3) and the number of mentions per account (Fig. 4). The first can give information of the accounts that are tweeting most actively. The three most productive accounts have tweeted more than 2000 tweets about climate: Michel de Muelenaere, a journalist working for *Le Soir*, Enrico Balducci, a public health professional, and Info climat, an information account about climate. Interestingly, none of them are among the twenty most mentioned accounts. In this case, the three on the top of the list are all media or information accounts, @RTBFinfo, @lemondefr and @Youtube. Among the 10 most mentioned, there are several accounts from France media, which shows the pervasiveness of news content among the two countries. Besides the general information providers, several accounts specifically related to climate change or sustainability are also among the top ten, such as @Reporterre, @DamienERNST1, @Goodplanet_ or @Novethic, as well as the green party, @Ecolo.

¹ This group is not only composed of lay people, but some stakeholders are also included, like associations, so a further subdivision could certainly be made.

It is also important to note that 50% of the data comes from accounts that have 55 or less tweets, and that belong probably to anonymous Twitter users, whereas the most productive users behave more like communication professionals.

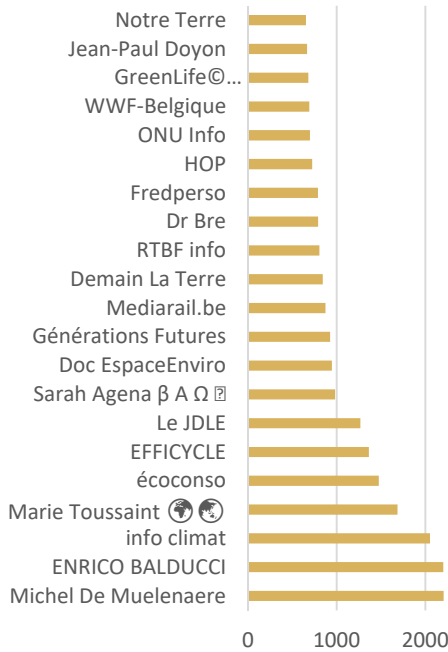


Figure 3 Number of tweets per account (top 20).

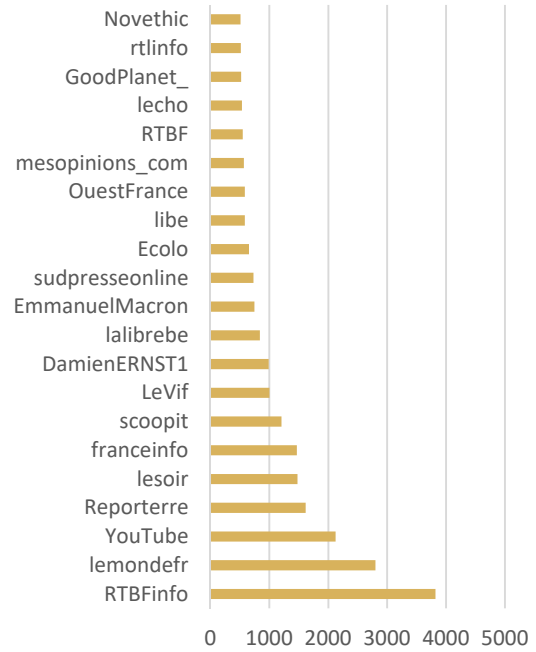


Figure 4 Number of mentions per account (top 20).

It is also relevant to look more in detail at the most productive accounts among the politicians (Fig. 5) and the media (Fig. 6). For the first, Ecolo is by far the most active party, which is not surprising considering ecology is one of their central objectives. In the case of the media accounts, the most productive are all well-known media: the first seven (as shown in the legend) are responsible for half of the tweets of their segment.

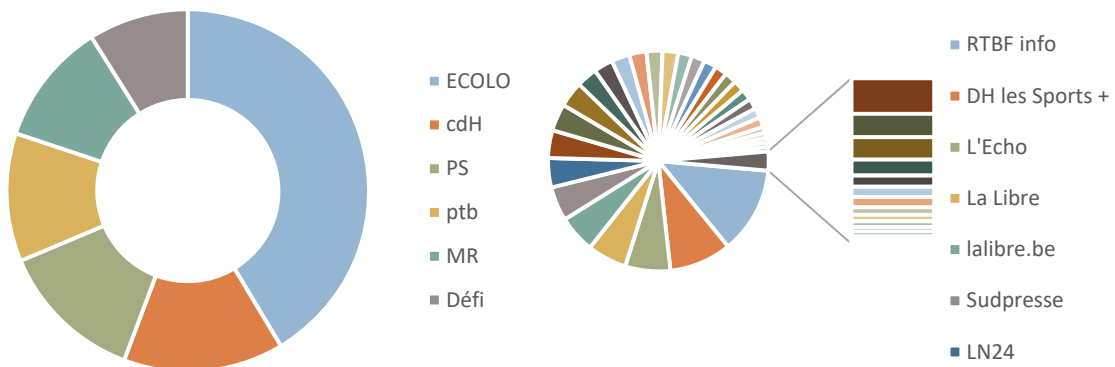


Figure 5 Tweets posted by political parties about climate change.

Figure 6 Tweets posted by media accounts about climate change.

It is also relevant to look in detail at which accounts are the most mentioned specifically by the politicians (Fig.7) and the population (Fig.8). First of all, only two accounts reach the top 10 for both segments: @RTBFinfo and @lesoir, which are leading media in the French speaking area (an audiovisual channel and a journal, respectively). For the rest, the crucial difference is that the population only mentions media channels (both Belgian and French, such as @lemondefr, which is in second positions) and @DamienERNST1, a university professor very active in the (social) media. On the other hand, the politicians mention mostly political accounts: from politicians, @RaulHedebouw (PVDA), @Marco_VanHees (PVDA), @CharlesMichel (MR), @jmnollet (Ecolo), from parties (@Ecolo, @lecdh, @MR_officiel) and official political institutions (@ParlWallonie).

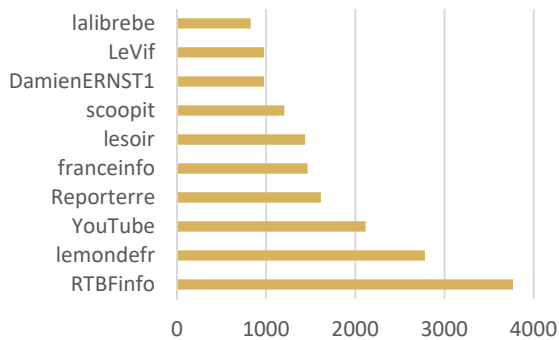


Figure 7 Most mentioned accounts by the population (Top 10).

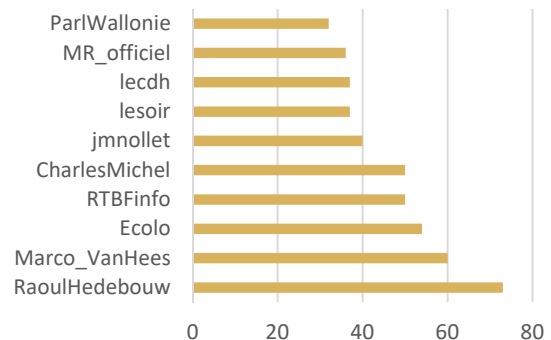


Figure 8 Most mentioned accounts by the politicians (Top 10).

When looking more in detail at what parties are responsible for the mentions of politicians and parties in the top 10 (Fig. 9), it appears that they are mostly (if not only) mentioned by their own, which is coherent with previous findings that show that Twitter is used for self-promotion by politicians (Coeseemans & De Cock 2017). The exception of @CharlesMichel and his party @MR_Officiel, for whom more than 40% of the mentions come from other parties. This is probably due to the prominent positions held by Charles Michel as Prime Minister (2014-2019) and President of the European Council (since 2019).

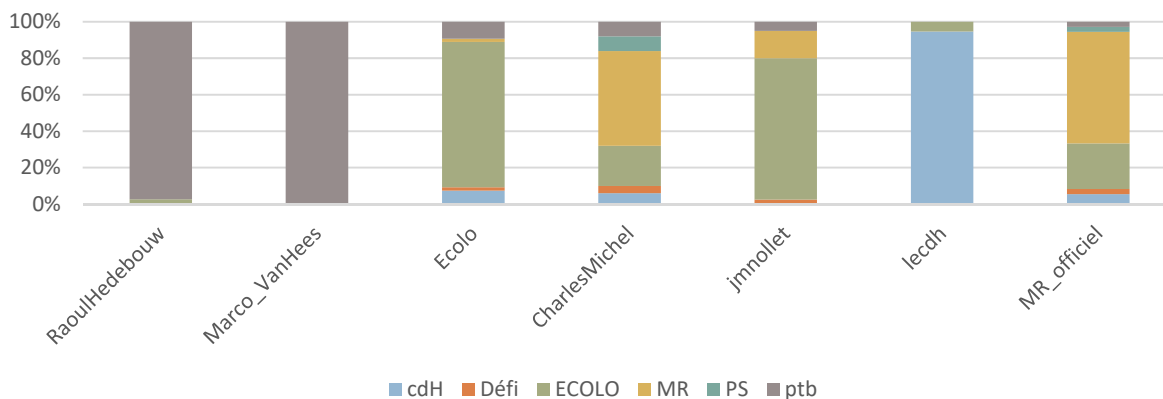


Figure 9 Mentions of political accounts by political parties (Top ten most mentioned accounts).

In the following section, we will describe the methods used in order to analyze the data from a linguistic perspective, in order to analyze our topics of interest (food, transport, energetic efficiency) and the expressions of difficulty.

2.2. ANALYTICAL METHODS: CORPUS LINGUISTICS

Our analysis uses a combination of quantitative and qualitative techniques, mostly from the subdiscipline of Corpus Linguistics. This groups several methods that allow to automatically analyze large amounts of data in order to get a quantitative overview of the trends in a corpus. For this particular report we have used descriptive statistics using software such as Excel, R and SketchEngine.

In what respects qualitative analysis, we have manually annotated the database with information about the axes under study (food, transport, energy efficiency) and the expression of difficulties. First, we have classified all the tweets according to whether they mention the topics of food, transport and energy efficiency or not. This is done through a semiautomatic technique based on the creation of an axe-related term list, and the application of each term as a filter in order to annotate the tweet as being about one of our three axes or not. The list was built based on the documentation from Ecoconso² (a well-known, Belgian information website about ecological habits) on each particular axe³.

Second, in order to analyze the expression of difficulty, we also created a list of French expressions typically used to express that concept. The list is rather eclectic, since difficulty is a meaning that can be conveyed through nouns, adjectives, adverbs, constructions, or emojis, a.o.

² <https://www.ecoconso.be/fr/content/climat-arrete-den-faire-des-tonnes>

³ Three researchers agreed on the configuration of the list: Louise-Amélie Cougnon, Vanessa Marcella and Andrea Pizarro Pedraza.

In order to build a comprehensive list, we started by writing an initial list of difficulty expressions⁴. We used that list as input for the Thesaurus tool in SketchEngine, which uses word vectors to calculate word similarity. The output gave us a new list of terms that were used in similar contexts in the corpus. With that, we created an extended list of potential difficulty expressions. Those were then checked in the corpus by three researchers for approval. At least two researchers checked every term. In case of doubt or disagreement, the term was rejected. With the final list, each difficulty term or expression was applied as a filter and annotated in the database as a new variable. Thus each tweet is classified as containing or not an expression of difficulty.

As for quantitative methods and corpus analytical tools, we have used SketchEngine for keywords (cf. 3.1.1.), key multiword expressions (3.1.2.), collocations (the inspection of the typical contexts of words) and similar expressions (through the Thesaurus tool, as explained in the previous paragraph.) We have created contingency tables and figures in Excel.

3. RESULTS

3.1. GENERAL TRENDS IN THE CORPUS

3.1.1. MOST TYPICAL TERMS OF THE CORPUS

One of the most revealing aspects to look at in a corpus are the keywords, that is the words that are more typical in the corpus under analysis than in a general corpus. They can give an idea of the dominant topic of a corpus. In this case, we have compared the BFCC with a random, general, Belgian French Twitter corpus as reference.

A keyword analysis is somewhat redundant in our case, considering our collection method. Our corpus was already built based on a list of expressions, which we obviously find among the first keywords (see Table 1). However, the keyword list is much broader than our original one and includes therefore many other terms. Moreover it does provide a quantitative view on which words are more typical (because they score higher).

⁴ Three researchers participated in this task: Louise-Amélie Cougnon, Marie Serisier and Andrea Pizarro Pedraza.

Table 1 List of keywords in the BTCCC as compared to a general Belgian French Twitter sample (20 first rows.)

Item	Frequency (focus)	Frequency (reference)	Relative frequency (focus)	Relative frequency (reference)	Score
sécheresse	3810	4	331.418	1.82239	117.779
biodiversite	1268	0	110.2987	0	111.299
recycler	6074	9	528.3551	4.10037	103.788
écologie	19104	35	1661.787	15.9459	98.123
durable	27249	51	2370.291	23.23545	97.844
sceptique	3305	5	287.4899	2.27799	88.008
recyclage	7758	16	674.8402	7.28955	81.529
ecocide	917	0	79.76649	0	80.766
déforestation	3856	7	335.4194	3.18918	80.307
climatechange	5533	11	481.2955	5.01157	80.228
photovoltaïque	3851	7	334.9845	3.18918	80.203
carbone	10270	23	893.3499	10.47873	77.914
environnement	41663	100	3624.113	45.5597	77.859
réchauffement	13399	31	1165.53	14.12351	77.134
dechets	846	0	73.59045	0	74.59
greta	6914	16	601.4237	7.28955	72.673
pollution	27052	70	2353.155	31.89179	71.573
climat	62647	167	5449.434	76.08471	70.707

Upon inspection of the first 100 terms, the keyword list gives us an idea of the many topics that are mentioned when tweeting about climate change: what climate change entails (*réchauffement, sécheresse*), what makes it worse (*écocide, déforestation*), what can help against it (*écologie, recyclage, durable, photovoltaïque*), how people feel about it (*sceptique*), who is important in the climate change debate (*Greta*), etc. We observe some terms that were not on our initial list that are also highly related to climate change, such as *anxiogène, ramassage* ou *atmosphérique*. The results also include variants of the lexical fields that were present in our list, such as *durabilité* ou *environmental* (from the fields of *durable* and *environment*).

It is important to mention that we find many hashtags (without the symbol # itself) among the keywords (for instance, *climatechange, marchepourleclimat, transitionénergétique..*) because they are part of the body of the tweet and are consequently taken into account for the keyword analysis. It would be possible to exclude them, but since they are sometimes embedded in the syntax of the tweet, we would lose information. Among them, many are in English (*upcycling, occupyforclimate, fridaysforfuture*) because climate change activism is international and their hashtags are very often used in their original language, and not translated.

3.1.2. MOST TYPICAL MULTIWORD EXPRESSIONS OF THE CORPUS

Keywords can also be calculated for multiword expressions, which gives back a list of typical word combinations that appear more often in our focus corpus than in the general one. These are expressions such as *développement durable*, *réchauffement climatique* or *écologie politique*.

Among the 20 first combinations (Table 2), we find meaningful groups of expressions that are based on the same adjective or noun. One of these groups is based on the adjective *climatique* (*réchauffement climatique*, *changement climatique*, *urgence climatique* and *défi climatique*). The nouns that are combined with it show conceptual differences in the way people refer to the climatic situation, focusing on the physical aspects (*réchauffement*, *changement*) or the social aspects (*urgence*, *défi*). Another group is built around the adjective *durable* (*développement durable*, *alimentation durable* or *mobilité durable*) that refer to changes of habits in different crucial aspects of people's lives. Another interesting group is built around the concept of *carbone* and contains recent expressions that name the quantified impact of people's habits on pollution: *taxe carbone*, *bilan carbone*, *neutralité carbone*, *empreinte carbone*.

Table 2 List of multiword expressions (key terms) in the BFCCC as compared to a general Belgian French Twitter sample (20 first rows.)

	Word	Focus	Reference	Focus	Reference	Score
1	développement durable	5,880	9	511.48	4.10	100.5
2	réchauffement climatique	11,633	20	1,011.91	9.11	100.2
3	écologie politique	967	0	84.12	0.00	85.1
4	taxe carbone	865	0	75.24	0.00	76.2
5	changement climatique	7,880	21	685.45	9.57	65.0
6	loi climat	1,402	2	121.95	0.91	64.3
7	transition énergétique	3,377	8	293.75	3.64	63.5
8	urgence climatique	2,297	5	199.81	2.28	61.3
9	alimentation durable	675	0	58.72	0.00	59.7
10	bilan carbone	846	1	73.59	0.46	51.2
11	greta thunberg	3,099	11	269.57	5.01	45.0
12	déchets nucléaires	1,385	4	120.48	1.82	43.0
13	déchets plastiques	897	2	78.03	0.91	41.3
14	paix durable	456	0	39.67	0.00	40.7
15	défi climatique	450	0	39.14	0.00	40.1
16	neutralité carbone	853	2	74.20	0.91	39.3
17	mobilité durable	646	1	56.19	0.46	39.3
18	crise climatique	1,257	4	109.34	1.82	39.1
19	transition écologique	3,706	16	322.37	7.29	39.0
20	empreinte carbone	1,028	3	89.42	1.37	38.2

3.2. TRENDS RELATED TO FOOD, TRANSPORT AND ENERGETIC EFFICIENCY

3.2.1. DEFINITION AND SCOPE OF THE THREE AXES

The BFCCC is very varied and includes all sorts of topics related to climate change, as we have mentioned in the previous section. In order to focus our analysis on parts of the corpus that might help us answer our research questions, we have decided to focus on three big topics that are intrinsically related to aspects of people's everyday lives that can affect climate change: food, transport and energetic efficiency. These three axes have been selected because they affect individual behaviors, as opposed to other aspects such as nuclear energy which are very much related to climate change, but are not in the hands of individuals. Taking action against climate change through food, transport and energy efficiency habits is potentially in people's control, and might thus be the source of actual obstacles. Understanding those difficulties is at heart of this study (see 3.3.) but first, we will describe the general trends related to these three axes.

Food production has a considerable impact on climate change that amounts to 30% of all emissions (Leiserowitz et al. 2020). However, some food habits have more impact than others. For instance, the 2013 FAO report about livestock-related emissions states that “greenhouse gas (GHG) emissions associated with livestock supply chains add up to 7.1 gigatonnes (GT) of carbon dioxide equivalent (CO₂-eq) per year – or 14.5 percent of all human-caused GHG releases” (Gerber et al. 2013). This is due to (an excessive) meat and dairy consumption in the world. Alternatives such as substituting red meat for chicken or fish, or ideally adopting a plant-based diet, are some of the recommendations to reduce the impact of food habits on climate change. Moreover, not all fruits and vegetables are equal: importing products from far away increases their carbon footprint, due to transportation emissions. Therefore, food not only involves eating but also shopping habits, namely buying local and with moderation.

In order to investigate this, we have annotated the corpus for food, beverages or food-related terms⁵ that are mentioned when speaking of eating habits that have a positive or negative impact on climate change. Some of these terms refer to foods or beverages whose production has a negative impact on climate change (*viande, fromage, bière, ...*) or their alternatives (*fruit, légume, soja,...*), to the negative or doubtful production methods themselves (*pesticides, OGM*), while others refer to eating or food-related habits that have a positive impact on climate change (*compost, manger bio/durable/local/moins, périmé...*).

According to the EU Climate Action⁶, our second axe of interest, transport, “represents almost a quarter of Europe's greenhouse gas emissions and is the main cause of air pollution in cities.” In

⁵ List of terms: *bière, burger, compost, épiluchure, fast food, fromage, fruit, lait, légume, manger bio, manger durable, manger local, manger moins, OGM, péremption, périmé, pesticides, poisson, porc, poulet, soja, steak, sucre, sulfites, tofu, viande, vin.*

⁶ https://ec.europa.eu/clima/eu-action/transport-emissions_en

this context, road transport is by far the most polluting transport, followed by civil aviation and navigation. Actions related to the use of cars are therefore prominent in the public debate: using alternative transportation means such as public transport or bikes, buying an electric car, or avoiding driving alone in rush hours.

We have annotated our corpus for terms referring to transports that people may use at an individual level⁷. Since our research focuses on personal obstacles, we did not include terms related to commercial transportation. Some of our terms refer to polluting vehicles (*avion, voiture*), to non- or less-polluting vehicles (*voiture électrique, vélo, transports en commun, trottinette...*), to alternatives such as vehicle sharing (*co-voiturage, autolib, velib*), to activities that could involve a change in transport use (*télétravail, déplacements professionnels, tourisme...*) and the Swedish loanword referring to the shame of flying (*flygskam*).

Our third topic, energy efficiency, is related to the impact of energy consumption. Despite the increase of renewable energies in Europe, in 2015, 72,6% of energy consumption was issued of fossil energy sources⁸, which are very polluting. Some countries, like Belgium, give financial help to make houses more efficient, by isolating or installing solar panels, green roofs, or a new heating system, for instance.

We have annotated the corpus with 25 terms related to energy efficiency in the household⁹. They refer to energy-consuming elements (*chauffage, chaudière, cuisinière à gaz*), alternative sources of energy or resources (*panneaux solaires, électricité verte, ...*) and energy-saving habits, actions or elements (*isoler, ampoule éco, citerne, ...*).

3.2.2. GENERAL DISTRIBUTION

In total, 26 047 tweets mention one of the annotated terms for the three axes, that is 6.7% of the global corpus. Among the three topics, transport and food are the most frequent, with 47% and 44% of these tweets respectively, whereas energy efficiency is only present in 9% (Fig. 10).

⁷ List of terms: *voiture, avion, vélo, tourisme, voiture électrique, train, bus, transports en commun, télétravail, à pied, tram, co-voiturage, trottinette, flygskam, batteries lithium, voiture hybride, autolib, marche à pied, déplacements professionnels, éco-tourisme, cambio, velib*.

⁸ <https://www.eea.europa.eu/fr/signaux/signaux-2017-1/articles/l2019energie-en-europe-situation-actuelle>

⁹ List of terms: *chauffage, rénovation, panneaux solaires, isoler, électricité verte, chaudière, ampoule éco, pompe à chaleur, châssis, douche courte, citerne, éteindre lumières, self-sufficiency, réduire eau, tiny house, récupérer eau, chauffe-eau solaire, thermostat, chauffer moins, chauffe-eau thermodynamique, pommeau de douche, double vitrage, cuisinière au gaz, pommeau de douche..*

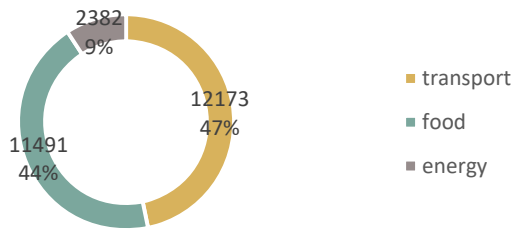


Figure 10 Overall distribution of tweets mentioning one of the three axes (in percentages.)

When observing the distribution of these axes across the social segments (Fig. 11), we notice that transport is equally mentioned by all the segments. For the other two, the percentages are roughly the same for the population and the media, whereas politicians show a greater interest in energy efficiency and less in food. This is particularly interesting since energy efficiency is overall the least mentioned topic, which reflects the relative importance of energy in political discourse while the population and the media are more concerned about food.

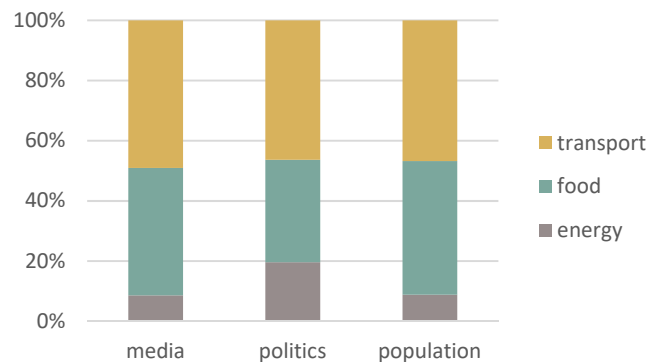


Figure 11 Distribution of tweets mentioning one of the three axes across segments (in percentages.)

In what respects the evolution of these topics over time (Fig. 12), transport was the dominant topic in the earlier years, while food steadily became more frequent until 2016, when it reached almost 65% of the data. In the most recent years, tweets mentioning transport have increased again while food has decreased, but has remained almost equally frequent. Energy, the least frequent topic, showed an increase in mentions around the year 2013 and then dropped continuously until 2016 where it reached its lowest point. Since then it has progressively reached 10% in 2021.

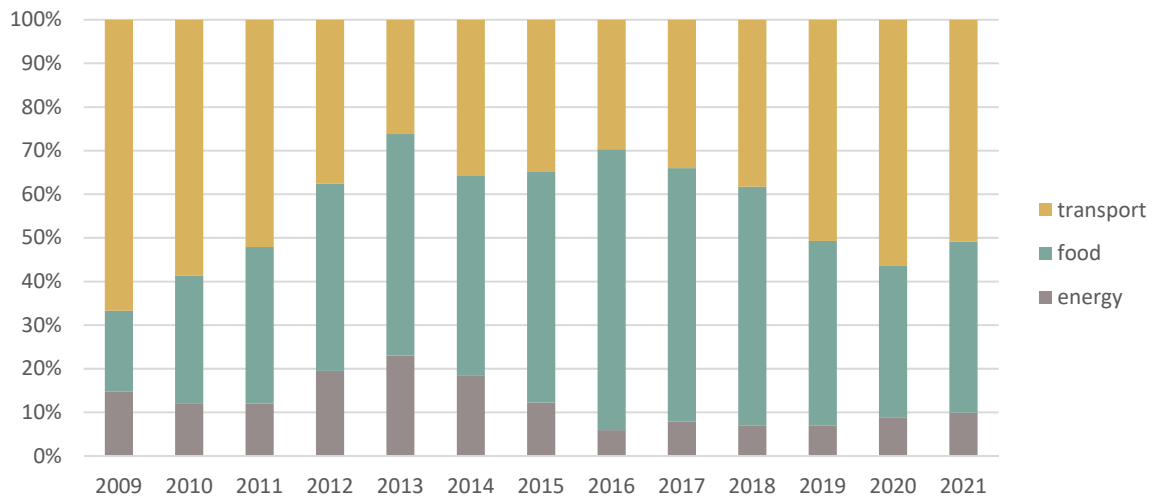


Figure 12 Distribution of tweets mentioning one of the three axes through time (in percentages.)

In the following sections, we will focus individually on the each of the axes.

3.2.1. TRENDS FOR FOOD TERMS

In total, food-related terms are mentioned 11 645 times: 10 661 individually (Figure 13) and 984 in combination (Fig. 14, top 10 combinations). The two most frequent terms (*pesticides*, *compost*) amount to half of the data, which indicates their relevance in this corpus. On top of the single-mention list, we find animal products (*viande*, *poisson*, *lait*), plant-based foods (*fruits*, *soja*) and *bière*, a term that is possibly so frequently mentioned due to its relevance in Belgium. Moreover we find frequent mentions to concepts related to foods such as avoiding waste (*zéro déchet*) and production methods that are subject to debate (*OGM*).

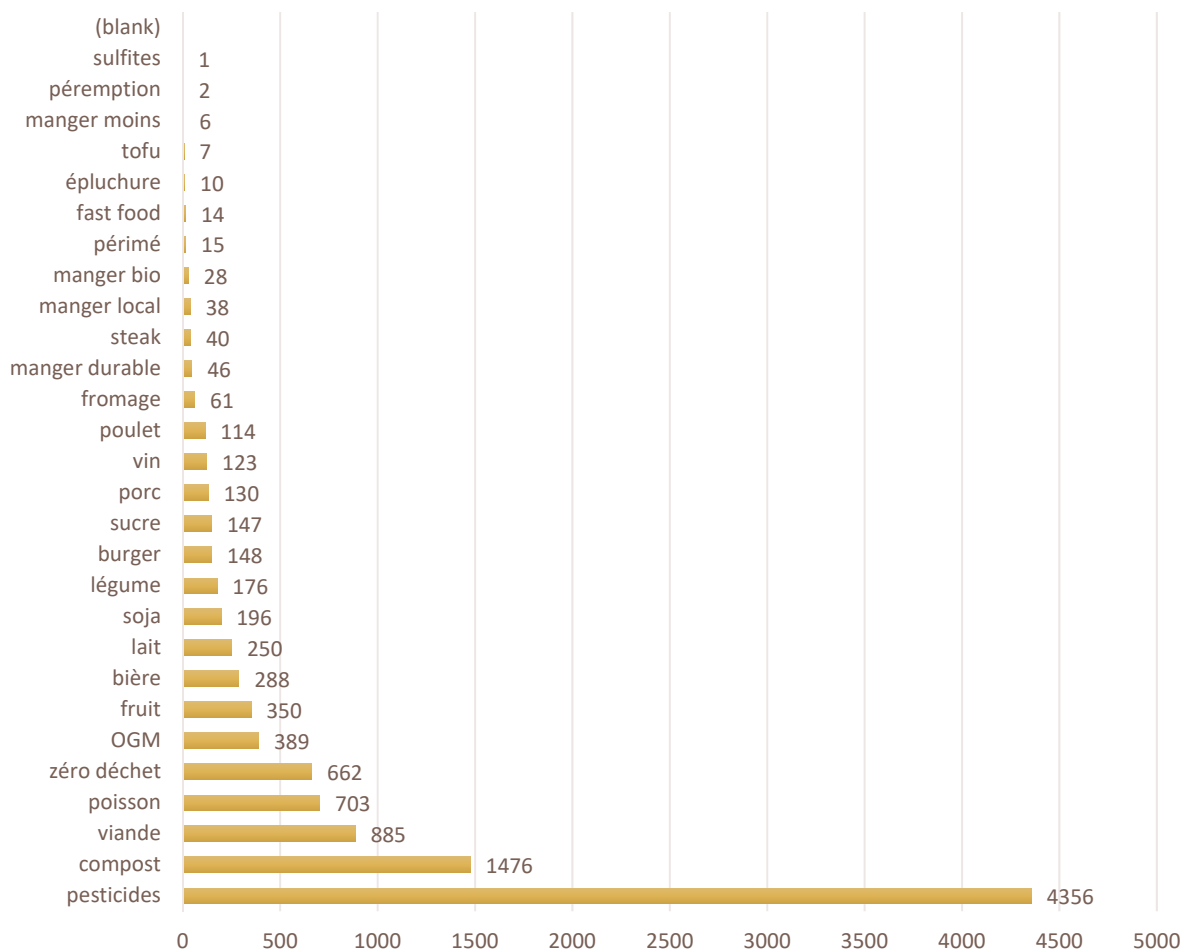


Figure 13 Number of tweets mentioning food-related terms (in raw numbers, not including combinations of 2 or more terms.)

The combinations of some of these terms are actually quite frequent and reveal what concepts are often associated in one single tweet. The most frequent association (*OGM*, *pesticides*) possibly points to the discussion of whether genetically modified organisms could reduce the use of pesticides. Several of these combinations include the term *viande*: some include another animal product, such as *viande/lait*, *viande/burger* and *viande/poisson*, and others include a plant-based

alternative (*viande/soja*). Interestingly, the concept is also often used in combination with *voiture*, which indicates a relevant association of these two aspects from different axes as having an impact on climate change at the level of the individual.

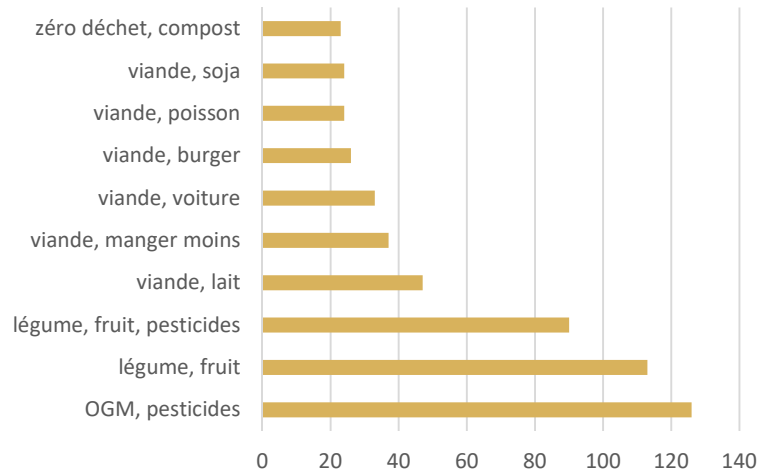


Figure 14 Tweets mentioning at least 2 food-related terms (10 most frequent combinations, in raw numbers.)

When looking at the evolution of food mentions through time (especially from 2012, when we have more than 100 data points) (Fig. 15) we observe a fluctuation in the most frequent term (*pesticides*) which reached its maximum percentage in 2013 and 2016, and remains somewhat constant around 40% in recent years. Some have slightly increased (*viande, lait, OGM, fruit, soja*) and others have had the opposite evolution (*zéro déchet*).

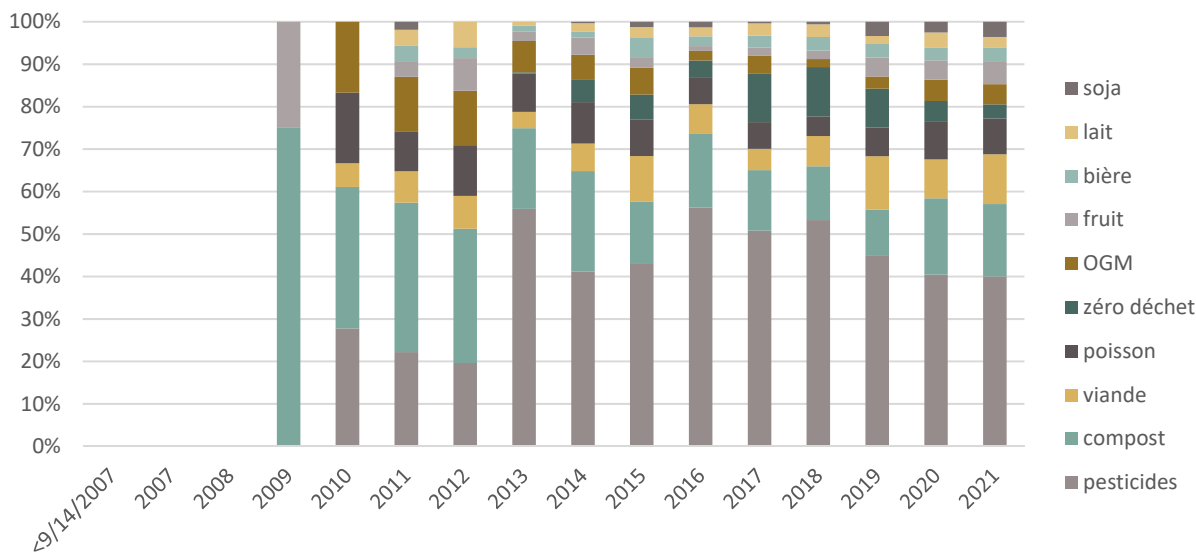


Figure 15 Distribution of tweets mentioning foods through time (top 10 most frequently mentioned foods, in percentages.)

The distribution of tweets mentioning food terms across the three different segments (Fig. 16) shows similarities between the population and the politicians, which use the top ten terms roughly in the same proportions. Some minor differences are that politicians mention *pesticides*, *zéro déchet* and *soja* more, while the population tweets more about *compost*, *poisson* or *bière*, proportionally speaking. The media accounts mention the different food terms in a more balanced way: almost all of them amount to 10% to 20% of the total. That means that the terms *poisson*, *zéro déchet*, *fruit* and *bière* are mentioned proportionally more by media accounts than by the other two segments. However, their tendency to balance does not include *OGM*, *lait* and *soja* which are much less frequent and show values similar to those of the other two social segments.

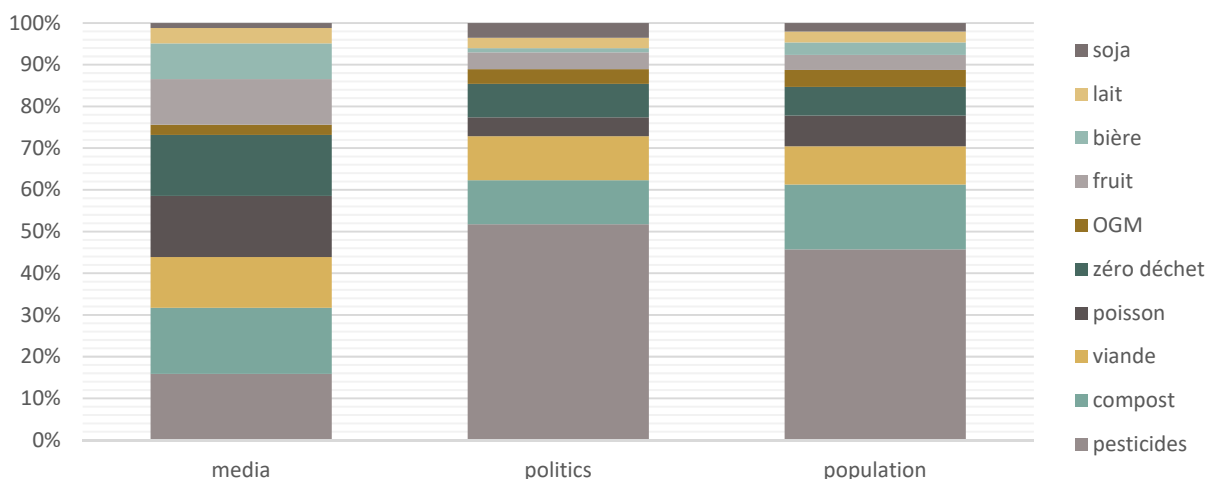


Figure 16 Distribution of tweets mentioning foods across segments (top 10 most frequently mentioned foods, in percentages.)

3.2.2. TRENDS FOR TRANSPORT TERMS

Transports are mentioned 12 204 times, out of which 11 515 are single mentions and 689 in combination. The term *voiture* is responsible for almost half of the data (Fig. 17). It is followed by far by several other vehicles such as *avion*, *vélo*, *voiture électrique*, *train*, *bus* and transportation means, *à pied*, and also by situations that have an impact on transportation, and therefore an effect on climate, such as *tourisme* and *télétravail*. There are also mentions of carpooling, *co-voiturage*, and other forms of vehicle sharing, through proper names that are relevant to the Belgian context like *cambio*, *autolib* and *velib*, but they are much less frequent.

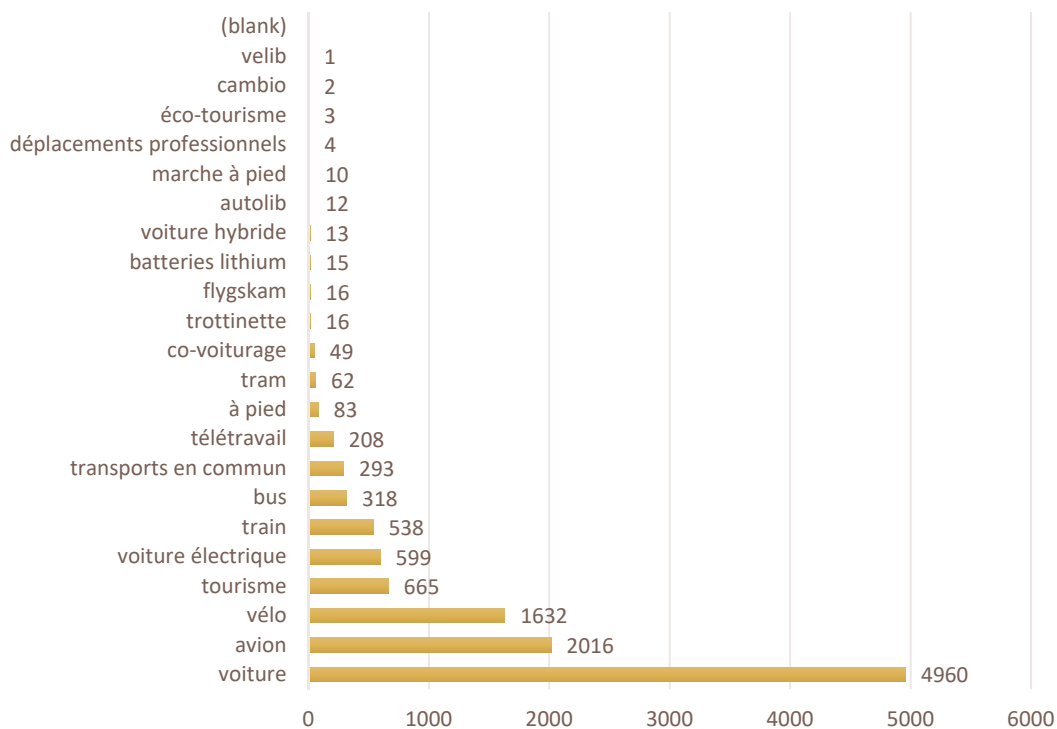


Figure 17 Number of tweets mentioning transport-related terms (in raw numbers, not including combinations of 2 or more terms.)

If we observe the combinations of transport terms in one single tweet (Fig. 18), we can see that the most common combinations are with the term *voiture*, either paired with alternative non- or less-polluting vehicles, like *voiture/vélo*, *voiture/transports en commun*, *voiture/train* or *voiture/bus*, or with the other most polluting transportation means for individuals, *avion/voiture*. Other combinations revolve around *vélo* in combination with other non- or less-polluting alternatives (*vélo/transports en commun* or *voiture électrique/vélo*) or with polluting vehicles (*avion/vélo*, *voiture/vélo/transports en commun*), possibly in order to discuss the alternatives.

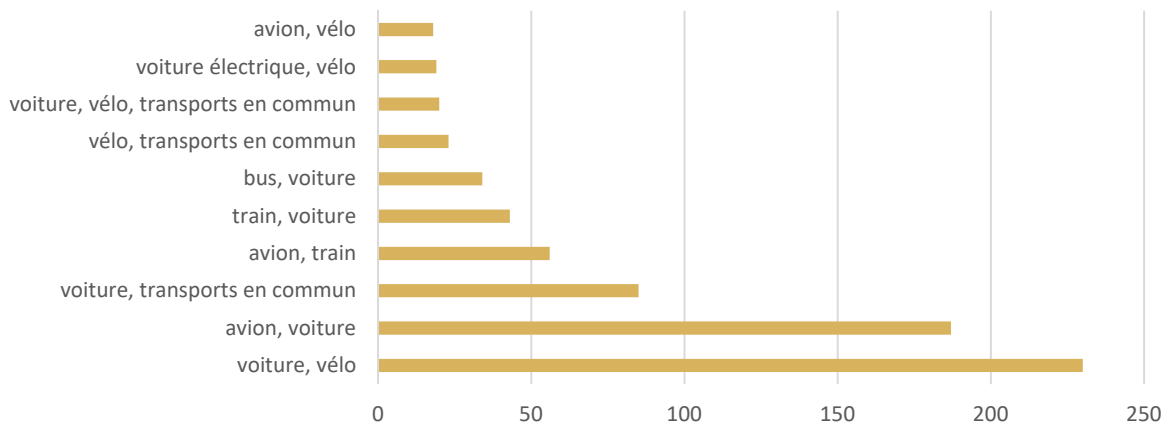


Figure 18 Tweets mentioning at least 2 transport-related terms (10 most frequent combinations, in raw numbers.)

In what respects the evolution of mentions through time (Fig. 19), the trends are less outspoken than for food terms. However, we still see that mentions for some terms have increased, such as *avion*, and others have decreased, like *tourisme*. For the rest, we rather observe similar proportions through the years with fluctuations that show no clear pattern. The tendencies shown by terms like *tourisme*, which seems to have dramatically decreased, should be taken with care, since the importance of the term might be boosted by the lack of data before 2012.

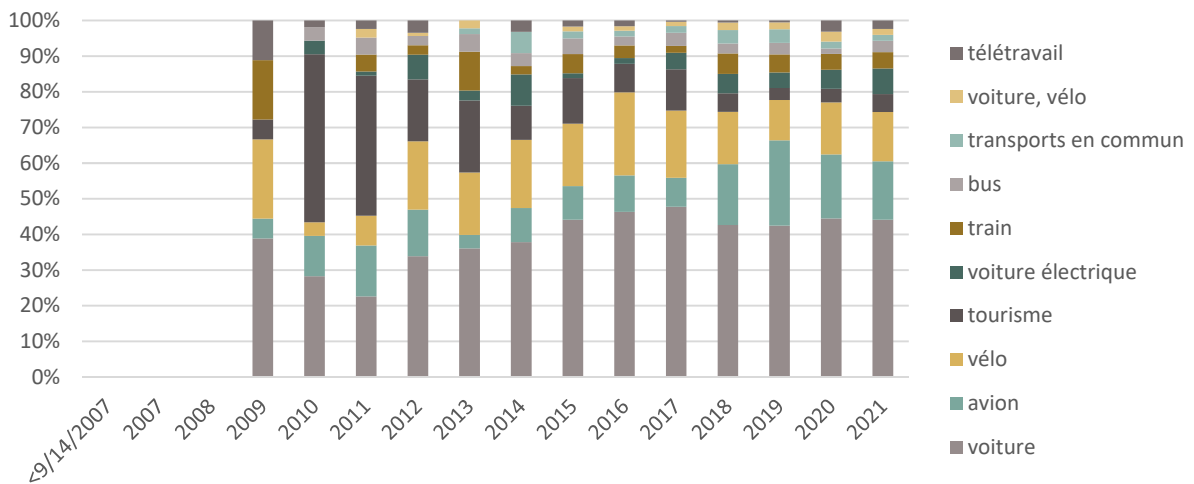


Figure 19 Distribution of tweets mentioning transports through time (top 10 most frequently mentioned, in percentages.)

When observing the distribution of transport terms across the three segments (Fig. 20), media and population accounts show similarities for the five most frequent terms (*voiture, avion, vélo, tourisme, voiture électrique*) which amount to almost 100% of the data for the media and almost 90% for the population. On the other hand, the politicians tweet less about the most polluting vehicles *voiture, avion*, and even *voiture électrique*, and more about *vélo* and *transports en*

commun. This is not surprising, since politicians want to encourage the population to make less use of the car, which is still the dominant and preferred transportation means of the Belgians. According to a federal survey about mobility¹⁰, 61% of all trips are done by car, while only 12 % are done by bike and 11 % by public transport.

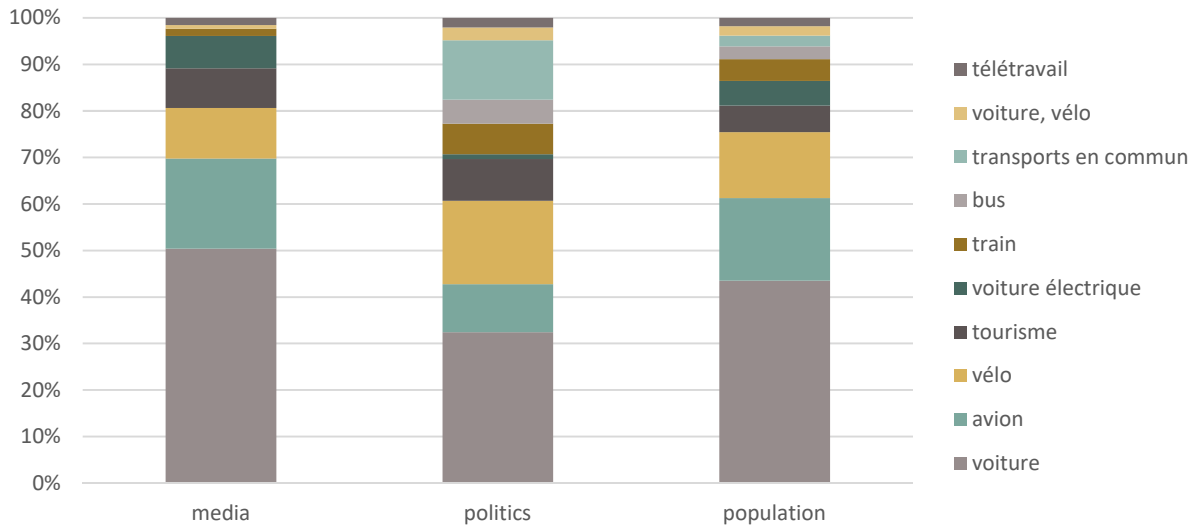


Figure 20 Distribution of tweets mentioning transports across segments (top 10 most frequently mentioned transports, in percentages.)

¹⁰ MONITOR : Enquête nationale sur la mobilité et la sécurité routière (2017), published in 2019.

3.2.3. TRENDS FOR ENERGETIC EFFICIENCY TERMS

In the corpus, energetic efficiency terms are the least frequent among the three axes under study. In total, the topic contains 2590 single terms and 135 combinations of terms. Only the two most frequent terms, *chauffage* and *renovation* account for almost 60% of the data. These are followed by far by some terms referring to renewable energy and its sources (*panneaux solaires*, *électricité verte*) and energy-saving methods (*isoler*, *chaudière*, *ampoule éco*, *pompe à chaleur*, *châssis*..)

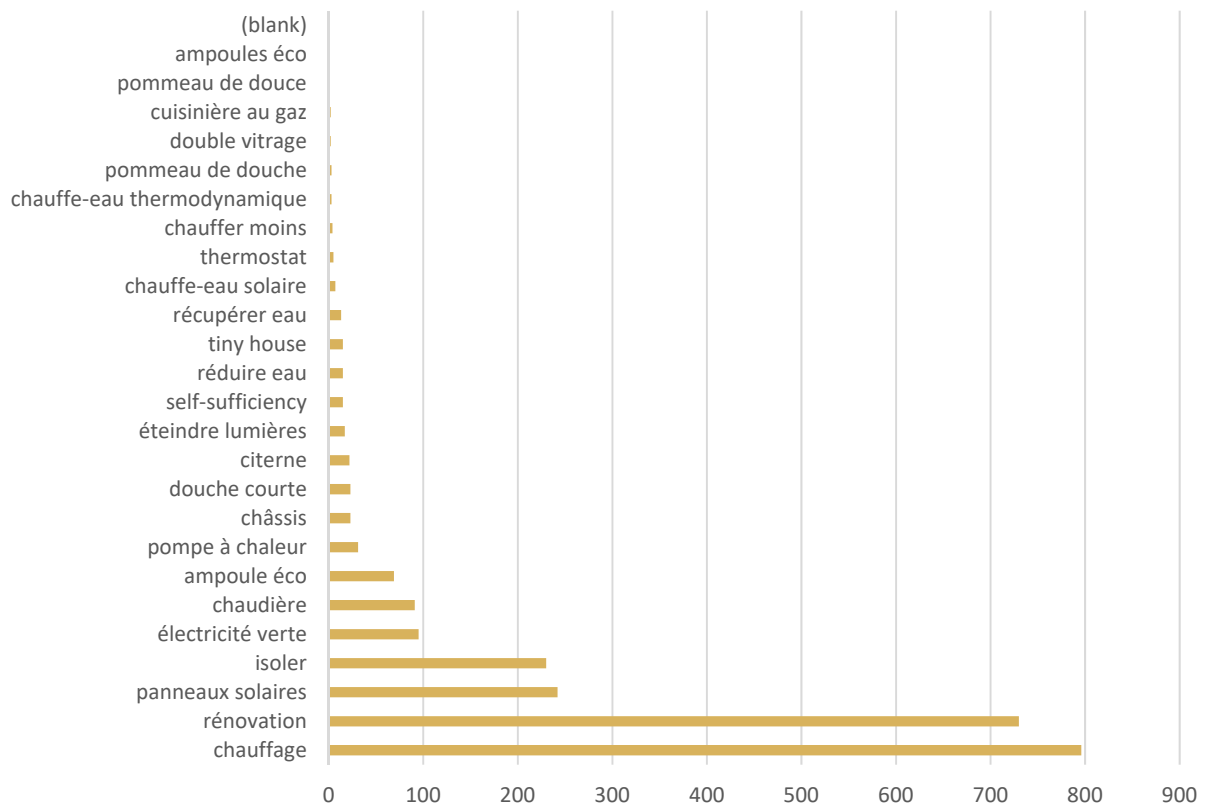


Figure 21 Number of tweets mentioning energetic efficiency-related terms (in raw numbers, not including combinations of 2 or more terms.)

The combinations of different terms in one single tweet point to associations around aspects of renovation (*isoler/rénovation*, *chauffage/rénovation*) and most of them around important home elements like the heating system (*chauffage/chaudière*) or the window frames (*châssis/panneaux solaires*, *châssis/rénovation*).

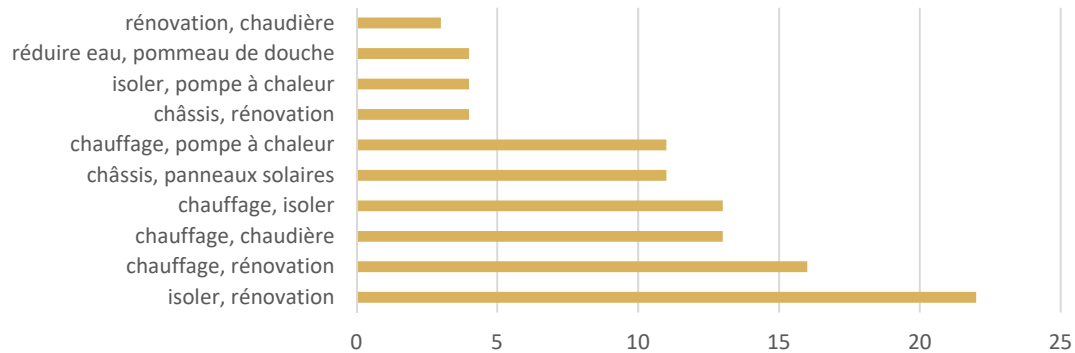


Figure 22 Tweets mentioning at least 2 energetic efficiency-related terms (10 most frequent combinations, in raw numbers.)

The distribution of the ten most common energy-efficiency terms through the years show no clear patterns, except for *chauffage* which used to be the most mentioned term until 2015 (although data before 2013 were very scarce) and then became the second most frequent after *rénovation*, which was mentioned more from then on.

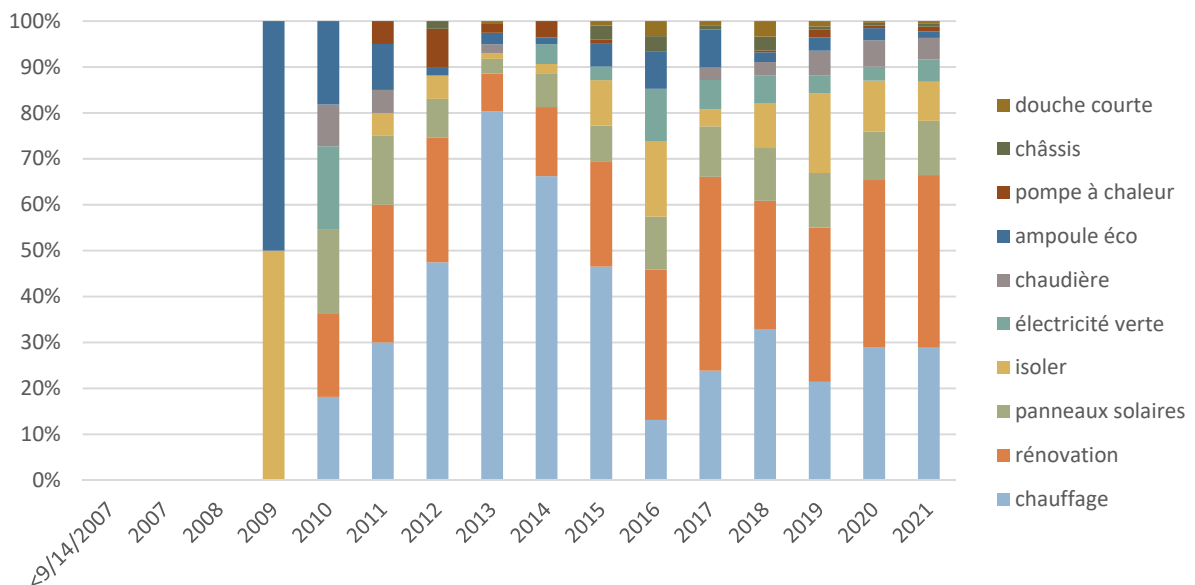


Figure 23 Distribution of tweets mentioning energetic efficiency through time (top 10 most frequently mentioned terms, in percentages.)

The distribution across segments shows that there are only 24 occurrences of energy-efficiency terms mentioned in media accounts, which is too few to reach any conclusions. For the other two segments, the most mentioned term by the population is *chauffage*, followed by *rénovation*, *panneaux solaires* and *isoler*, whereas the politicians tweet more about *rénovation*, followed by *isoler*, *chauffage*, and in third place *panneaux solaires* and *ampoules éco* with the same frequency.

Again, these preferences might point to the fact that politicians want to encourage renovation works in order to isolate private houses, for which the different Belgian regional governments allocate several bonuses. On the other hand, the population mentions heating more frequently, possibly because it affects their everyday lives and monthly expenses, and is therefore a more prominent problem.

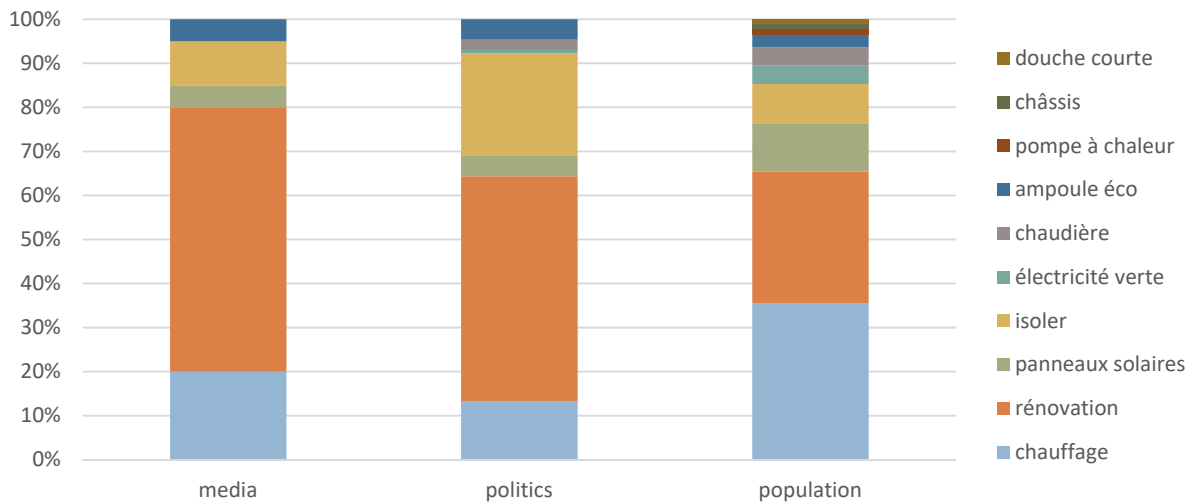


Figure 24 Distribution of tweets mentioning energetic efficiency across segments (top 10 most frequently mentioned terms, in percentages.)

3.3. THE EXPRESSION OF OBSTACLES IN TAKING ACTION AGAINST CLIMATE CHANGE

Finally, we wanted to explore the expression of difficulties and obstacles in taking action against climate change. As explained in 1.3, while climate change is widely accepted in Belgium, it is not easy for people to take action against it in everyday life. This analysis involved delimiting a list of linguistic expressions, at all linguistic levels, that conveyed difficulty in French. We take *difficulty* as a very large concept that expresses the ideas of *complication*, *problem*, and also *obstacle*. *Difficulty* is not a linguistic term per se like *negation* or *tense*; rather, it may be expressed through a variety of linguistic phenomena at different levels, such as nouns (*difficulté*), adjectives and their negations (*compliqué, pas simple*), constructions (*mais comment..., ne pas VERB comment*) and other pragmatic devices such as emojis (🤔 = thinking).

In total, 13 559 tweets included a difficulty expression (about 4% of the total corpus) (Fig. 25). The two most frequent expressions, *problème* and *défi*¹¹, account for 50% of the data.

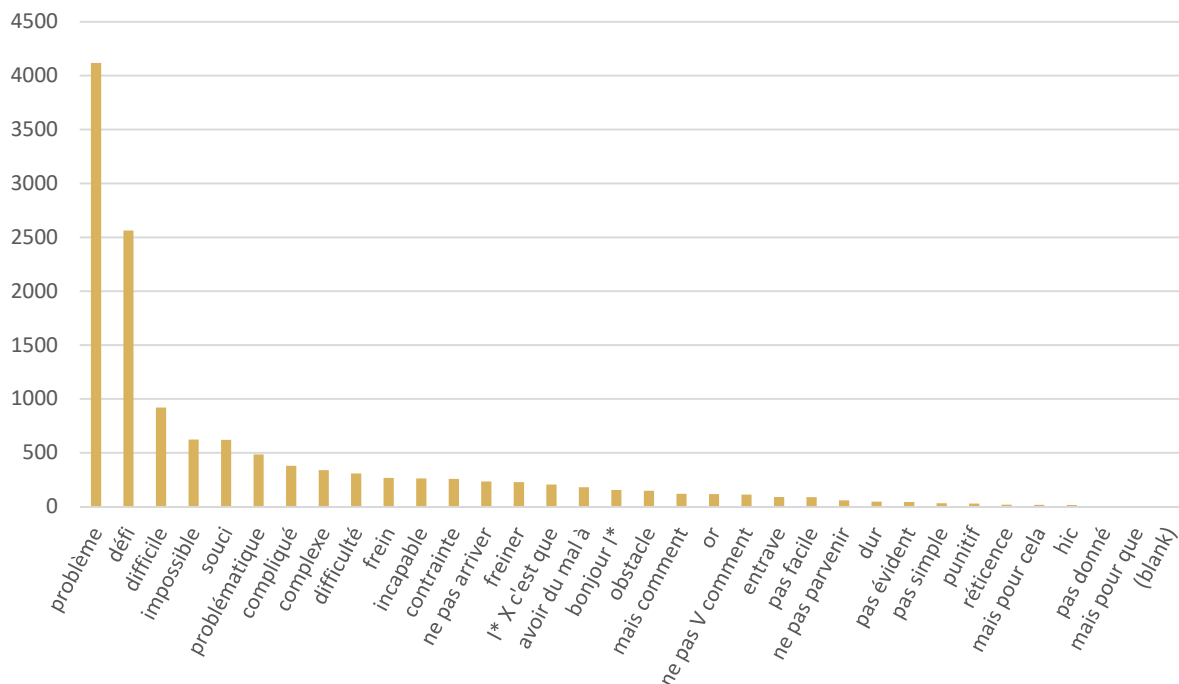


Figure 25 Expressions of difficulty (raw numbers, single use, excluding emojis.)

Out of the total, 556 tweets include a combination of different difficulty expression, for 143 different combinations. Fig 26 shows the 10 most frequent combinations.

¹¹ References to the Belgian political party Défi have been discarded from the data.

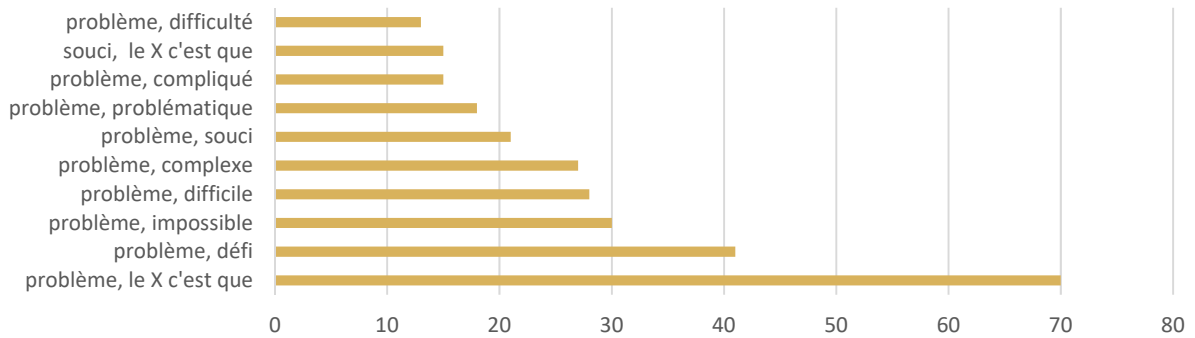


Figure 26 Expressions of difficulty in combination (raw numbers, top 10).

When looking at which segments express more difficulties, politicians are surprisingly the ones that use more expressions of difficulty, in 4.5% of their tweets (see Fig. 27).

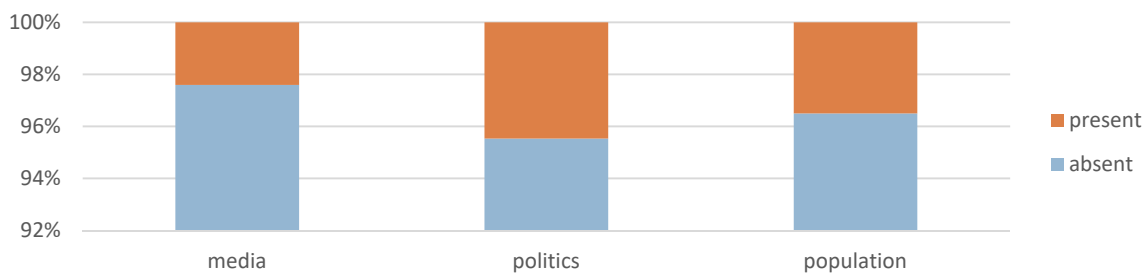


Figure 27 Presence/ Absence of an expression of difficulty in a tweet per segment (in percentages.)

When looking at the particular difficulty expressions used per segment (Fig. 28), overall, the media show less different types than the other two groups, although the apparently smaller variation might be due to the considerably smaller size of the subcorpus. The two most frequent terms, *problème* and *défi*, are not equally distributed across segments. We see that the use of *défi* (challenge) is preferred by media and mostly by politicians. The concept of *défi* includes in its semantics a potential overcoming of the problem, which is why it has a more positive connotation that might be more suitable for political discourse. In that same line of thought, *impossible* is entirely absent from the media and political discourse when speaking about climate. Therefore, although politicians use overall more difficulty expressions, the particular terms they use have less negative connotations.

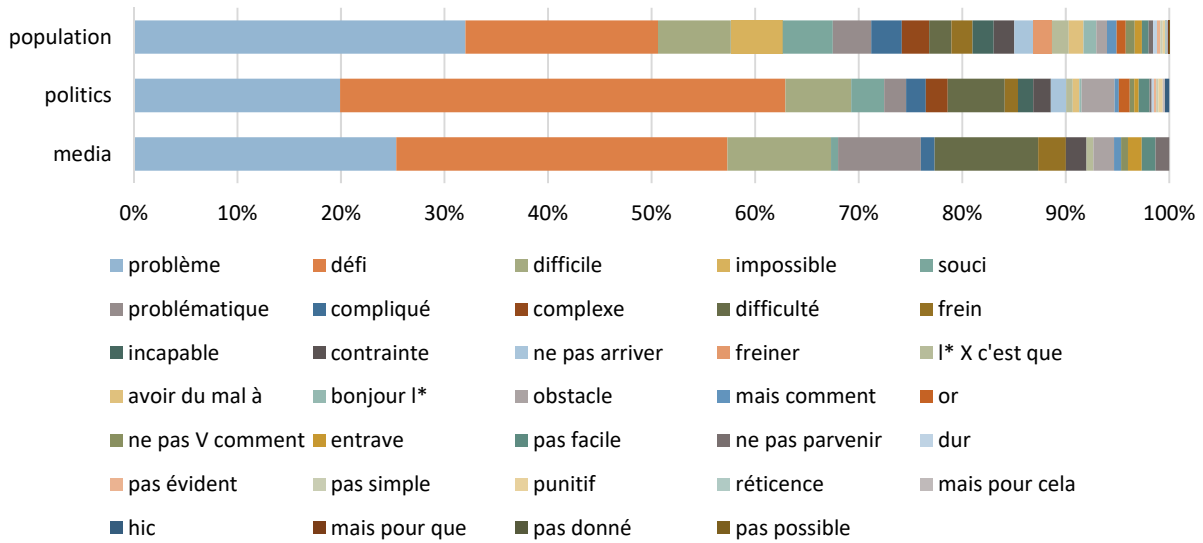


Figure 28 Distribution of particular expressions of difficulty per segment (in percentages)

As for the evolution over time (Fig. 29), although the percentages are small, there is a noticeable increase in the tweets that express difficulties from 2017 until 2021. This might indicate that there are more difficulties or that there is more awareness of the difficulties, or both. In fact, if we assume that, with time, the list of actions that citizens can take against climate change has itself increased, it is just normal that the difficulties in doing those actions have also multiplied. However, we can not discard the possibility that the population might also be more aware of the difficulties.

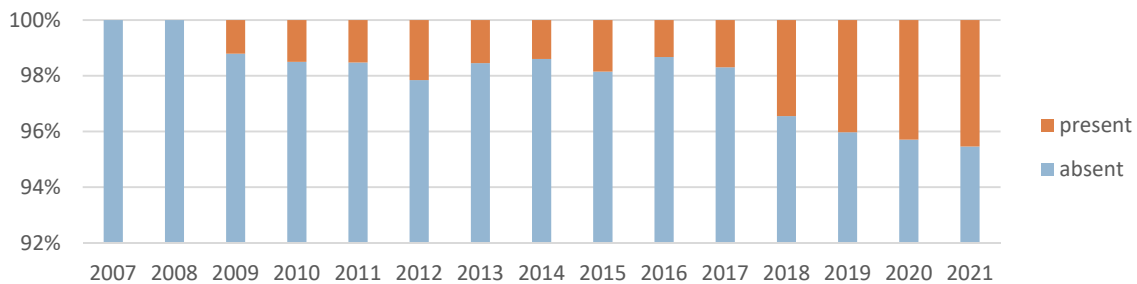


Figure 29 Presence/ Absence of an expression of difficulty in a tweet through time (in percentages).

In what respects the three axes (Fig. 30), there is also a difference for the topic of food, which is expressed in terms of difficulty less often (3%, versus more than 4% for the other two axes.) This indicates that actions against climate change in transport and energy efficiency present slightly more obstacles for individuals than those related to food.

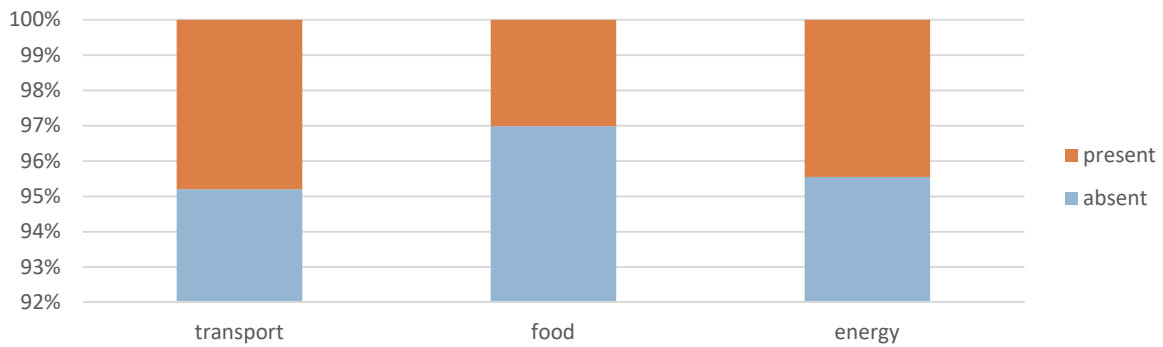


Figure 30 Presence/absence of an expression of difficulty in a tweet across axes (in percentages).

Finally, we have looked at emojis separately. Namely, we have selected those that convey a pragmatic meaning of difficulty and filtered the corpus to annotate their presence in a tweet, individually and in combination:

- 🤔: thinking is needed because something is difficult to understand
- 😕: confusion is the result of a difficulty
- 🙄 a raised eyebrow is provoked by a difficult question
- 🤷 shrugging shoulders is provoked by lack of knowledge/understanding/power
- 🤔 a difficult aspect requires careful consideration/pondering

When observing their distribution with the presence or absence of a difficulty expression (Fig. 31), it is noticeable that difficulty expressions are mostly combined with a single emoji, while emojis in combination do not appear in the same tweet as a difficulty expression. This might point to the fact that the accumulation of emojis serves the same purpose and no more linguistic expressions at other levels are needed.

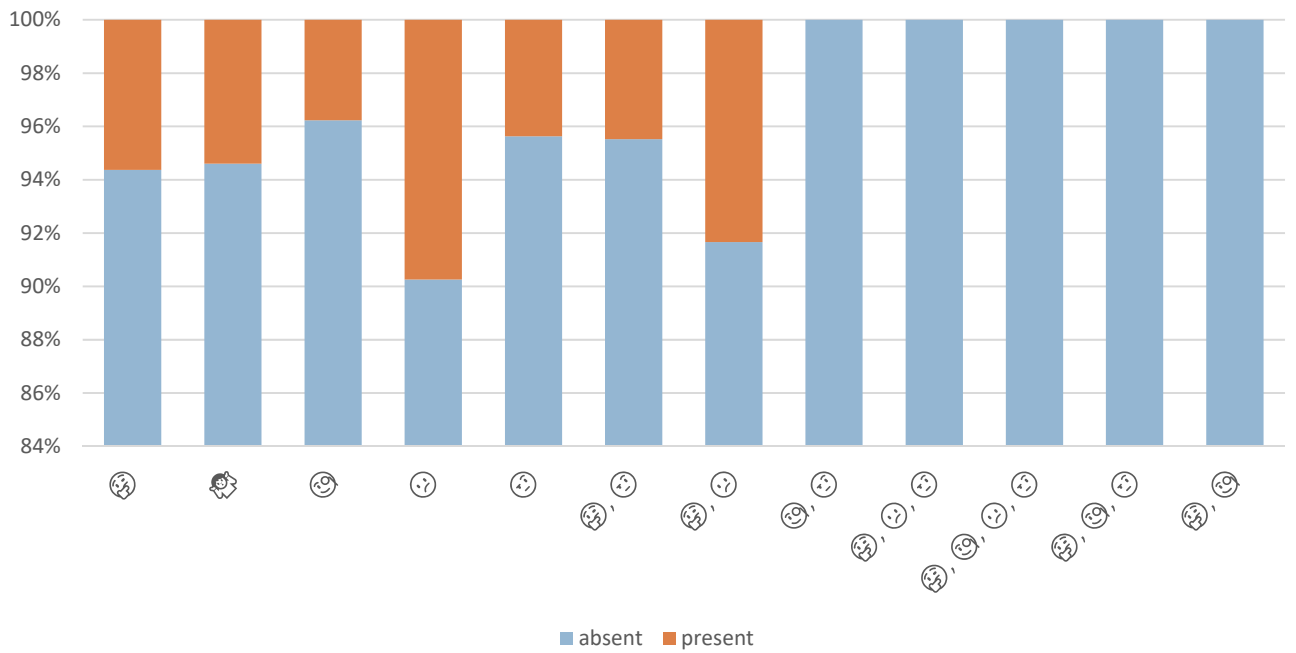


Figure 31 Presence/ Absence of an expression of difficulty in a tweet in combination with emojis (in percentages.)

Among the emojis, the one that is most frequently combined with a difficulty expression is the confused face 😬: almost 10% of the total occurrences of this emoji are in a tweet with a difficulty expression.

4. SUMMARY: KEY RESULTS OF THE REPORT

In this report, we have summarized the findings of the WP3 of the project “202CM *Overcoming Obstacles and Disincentives to Climate Change Mitigation: A cross-cutting approach by human and social sciences*”, funded by Solstice (JPI Climate). The initial goals were: 1. to collect a corpus of Belgian French, Climate change-related tweets; 2. to analyze the distribution of the data across segments, time and particular accounts (most active, most mentioned); 3. to analyze the presence and distribution of several important axes related to Climate change that have a crucial impact on individual life (food, transport, energetic efficiency) and 4. to extract and analyze the expressions of difficulty in the corpus at different linguistic levels.

For objective 1, we have successfully collected a Twitter corpus of Belgian French tweets related to Climate Change.

- The corpus has 385 977 tweets and it has metadata such as segment (media, population, politicians) or time.

As for objective 2, we have described the distribution of the corpus. Due to methodological reasons, the corpus is not equally distributed across segments and time: population accounts are responsible for most of the data (almost 370 000 tweets) and there is limited tweets for the earlier years.

- The most mentioned accounts are mostly from general media channels or journals, but also from some information or political accounts that are focused on sustainability and climate change.
- Politicians mostly mention political accounts from their own party for self-promotion, whereas the population only mentions media channels and influencers. Among the political parties, Ecolo is the party that tweets the most.
- The most typical keywords are *sécheresse*, *biodiversité* and *recycler*. Among the keywords, we also find hashtags in French and English, which reflect the international nature of the topic of climate change. The keyword analysis for multiword expressions gives back expressions around the terms *climatique*, *durable* and *carbone*. The varied combinations show conceptual differences in the way people refer to the climatic situation.

For objective 3, we annotated the corpus for terms related to food, transport and energetic efficiency, aspects through which people can have an individual impact on climate change. 6.7% of the corpus includes an expression of one of the axes.

- Transport and food are the most frequent, while energy efficiency only represents 9% of the axe-related tweets. Transport is mentioned equally by all segments, whereas energy efficiency is more mentioned by politicians, and food by the population. The mentions of food increased around 2016 and then decreased, while transport had the opposite evolution.
- The most frequent food-related terms are *pesticides* and *compost*. The frequency of use of certain terms has increased through the years (*viande*, *lait*, *OGM*, *fruit*, *soja*), on the other hand *zéro déchet* has decreased. The distribution of these terms across the three segments

is roughly the same for the population and the politicians, while the media mention the different food terms without clear preferences.

- Among the transport terms, the most mentioned is *voiture*. The fluctuations through time show no clear pattern. In this case, media and population accounts show similar preferences in the use of the most frequent terms, whereas the politicians avoid mentioning polluting vehicles (*voiture, avion*) and tweet more about alternatives (*vélo, transports en commun*).
- The most frequent terms for energetic efficiency are *chauffage* and *renovation*. In recent years, *renovation* has become the most frequently mentioned. It is also the most frequent term for the politicians, while the population tweets more about *chauffage*, showing that it is a more prominent concern for this segment.

Finally, objective 4 was to extract the expressions of difficulty so as to understand what are the obstacles for taking action against climate change. 31 types of expressions and 5 emojis were annotated (nouns, adjectifs, constructions), as well as their combinations. 13 559 tweets included a difficulty expression (about 4% of the total corpus).

- The two most frequent expressions are *problème* and *défi*.
- Politicians express difficulty proportionally more than the other segments, however, they prefer the term *défi* which implies that the difficulty can be overcome, and they avoid the adjective *impossible*.
- In general, tweets that express difficulties have increased through time, which might indicate that people experience more difficulties in taking action against climate change, or that there is more awareness of the difficulties. As for the axes, Twitter users express more difficulties while speaking about transport and energy efficiency than about food.
- Difficulty expressions are not often combined with more than one emoji at once, which indicates that their accumulation might be redundant. The emoji that is most frequently combined with an expression of difficulty is the confused face 😕.

5. REFERENCES

- Bouman, T., Steg, L., & Perlaviciute, G. (2021). From values to climate action. *Current Opinion in Psychology*. <https://doi.org/10.1016/j.copsyc.2021.04.010>
- Bourg, D., Grandjean, A., & Libaert, Th. (2006). *Environnement et entreprises : en finir avec les discours*, Pearson Education, Paris.
- Coesemans, Roel ; De Cock, Barbara. (2017) Self-reference by politicians on Twitter: Strategies to adapt to 140 characters. In: *Journal of Pragmatics*, Vol. 116, p. 37-50 <http://hdl.handle.net/2078.1/176070> -- DOI : 10.1016/j.pragma.2016.12.005
- Cougnon, L.-A. & de Viron, L. (2021, in press). Covid-19 et communication de crise. Focus linguistique sur les tweets francophones de Belgique. In Cotelli Kureth, S, Dubois, M. & Kamber, A. (éds). *La linguistique appliquée à l'ère digitale*, Bulletin suisse de linguistique appliquée,
- Gerber, P.J., Steinfeld, H., Henderson, B., Mottet, A., Opio, C., Dijkman, J., Falcucci, A. & Tempio, G. 2013. Tackling climate change through livestock – A global assessment of emissions and mitigation opportunities. Food and Agriculture Organization of the United Nations (FAO), Rome.
- Grundmann, R., and R. Krishnamurthy. 2010. "The discourse of climate change A corpus-based approach." *Critical Approaches to Discourse Analysis across Disciplines* 4 (2): 125–146.
- Grundmann, R., and M. Scott. 2014. "Disputed climate science in the media: Do countries matter?" *Public Understanding of Science* 23 (2): 220–235.
- Ifert Johnson, D. Roloff, M. E. & M. A. Riffée 2004 Politeness theory and refusals of requests: Face threat as a function of expressed obstacles, *Communication Studies*, 55:2, 227-238, DOI: 10.1080/10510970409388616
- Jang, S. M., and P. S. Hart. 2015. "Polarized frames on 'climate change' and 'global warming' across countries and states: Evidence from Twitter big data." *Global Environmental Change* 32: 11–17.
- Kirilenko, A. P., and S. O. Stepchenkova. 2014. "Public microblogging on climate change: One year of Twitter worldwide." *Global Environmental Change* 26: 171–182.
- Koteyko, N. 2010. "Mining the internet for linguistic and social data: An analysis of carbon compounds' in web feeds." *Discourse & Society* 21 (6): 655–674.
- Leiserowitz, A., Maibach, E., Rosenthal, S., Kotcher, J., Bergquist, P., Ballew, M., Goldberg, M., Gustafson, A., & Wang, X. (2020). *Climate Change in the American Mind: April 2020*. Yale University and George Mason University. New Haven, CT: Yale Program on Climate Change Communication

Mayer, F. W. 2012. "Stories of climate change: Competing narratives, the media, and U.S. public opinion 2001–2010." Joan Shorenstein Center on the Press, Politics and Public Policy Discussion Paper Series #D-72. https://www.hks.harvard.edu/presspol/publications/papers/discussion_papers/d72_mayer.pdf

Merner, C. (2017). *The Psychology of Effective Climate Change Communication: A Discourse Analysis*. PhD thesis, Dalhousie University, Canada.

Olteanu, A., C. Castillo, N. Diakopoulos, and K. Aberer. 2015. "Comparing events coverage in online news and social media: The case of climate change." Proceedings of the International AAAI Conference on Web and Social Media (ICWSM 2015).

Pearce, W., K. Holmberg, I. Hellsten, and B. Nerlich. 2014. "Climate change on Twitter: Topics, communities and conversations about the 2013 IPCC Working Group 1 Report." *PLoS ONE* 9 (4): e94785. doi: 10.1371/journal.pone.0094785

Salway, A. (2017) Data-Driven Approaches to Climate Change Discourse, Illustrated through Case Studies of Blogs and International Climate Negotiation. In Fløttum, K, ed. *The role of language in the climate change debate*. New York: Routledge.

Tvinnereim, E., and K. Fløttum. 2015. "Explaining topic prevalence in answers to open-ended survey questions about climate change." *Nature Climate Change* 5: 744–747.

Wild, K., A. Church, D. McCarthy, and J. Burgess. 2013. "Quantifying lexical usage: Vocabulary pertaining to ecosystems and the environment." *Corpora* 8 (1): 53–79.