UNIVERSITY OF PADOVA

Department of General Psychology*

Bachelor's Degree Course in Psychological Science

Final dissertation

Words during the pandemic:
a quantitative analysis of politicians' speeches in France

Supervisor
Professor Anna Spagnolli

*Co-supervisor (if present)***
Mariavittoria Masotina

Candidate: Olivia Bettella
Student ID number: 1217001

Academic Year 2021/2022

Words during the pandemic: a quantitative analysis of politicians' speeches in France

Words during the pandemic: a quantitative analysis of politicians' speeches in France

1. INTRODUCTION

The ways, or in this case *words*, with which prominent political figures all across the world addressed and communicated the events of the pandemic, shaped people's perception of risk and ultimately impacted the effective implementation of responses to the crisis. In this thesis, I will analyze the recurrence of specific terms across a selection of the most important speeches delivered during the first two years of the Covid-19 pandemic in France (January 2020 - December 2021) by two opposing French politicians. Both informants carefully utilized language in order to communicate with the public: one did so to justify suspending some of their otherwise uninfringeable freedoms, the other often to criticize those very demands implemented by the former.

1.1 Covid-19 in France: how it started

The first Covid-19 cases in France were reported as early as December 27th 2019 (Peillon, 2020), yet the first big outbreak was linked to a sizable religious gathering that took place in Mulhouse, in the Grand-Est region, between February 17th and 24th (Noël, 2020). Shortly thereafter, Grand-Est and the nearby Île de France (Paris and the surrounding areas) became epicenters of the pandemic in France.

Prior to March 16th - when French daily life shifted from having relatively light restrictions to a complete lockdown - the measures introduced at the time were mostly general recommendations for safe practices to stop the virus's transmission, these were frequently and widely disregarded. According to research from Desson et al. (2020), the administration chose to allow municipal elections to go forward as planned on March 15th despite non-essential shops closing on March 14th and the first official lockdown starting just two days later, on March 16th. The only changes made to voting procedures were priority lines for vulnerable people and social distancing recommendations. This inconsistent approach to policy extended all the way to border control measures, with France going from advising new arrivals to isolate themselves for 14 days, to a complete closure of their borders to non-essential travelers and the announcement of a nationwide lockdown just two days later. The combination of these things may have led to confusion amongst citizens regarding the severity of the disease.

On March 16th, when president Macron announced the first nationwide lockdown, one of his main opposers, founder and head of the left-wing populist party "La France Insoumisme" Jean-

Luc Mélenchon, uploaded a short message to his YouTube channel. In his speech (ME2_16-03-20), Mélenchon highlighted the confusion caused by the head of state up to that point and, although he declared his support for the decision to finally instate a national stay-at-home order, he still criticized the president for being too lenient: "… you can see, as I can, all the flaws in the government's plan for several days. Let's say that the weekend was especially calamitous. Terrible announcements on Saturday, but the election taking place the next day and now again, we still feel a kind of vagueness in the president's words. For example, he says you must stay home, but you can go out and get some exercise. No, no, you have to confine yourself, that's the rule now".

In spring of 2020, the global landscape was defined by extreme panic about the Covid-19 pandemic. On April 14th 2020, an article was published on the popular American news and opinion website "Vox" under the provocative title "How President Emmanuel Macron bungled France's coronavirus response" (Ward, 2020). In the article, the author explains how, despite France being one of Europe's most powerful countries with a world-renowned healthcare system and ample social welfare which could have provided a successful response to the health crisis, the mishandling of the pandemic could largely be attributed to the government of French President Emmanuel Macron and their inadequate response and crucial errors during the early stages of the Covid-19 epidemic.

Ever since Covid-19 first became a global public-health concern in January of 2020, France and its uniquely centralized presidential system were under public scrutiny. From minor news outlets to prominent state leaders, critique of the president's handling of the pandemic was widespread, making the French response, and the speeches surrounding it, a particularly interesting case analysis.

1.2 Overarching themes

This thesis allocates itself within a much larger research project which includes similar analyses of speeches made by government majorities and their opposition in several different countries during the Covid-19 pandemic. The research is being conducted by a group of psychology students from the university of Padova under the supervision of Prof. Anna Spagnolli and graduate student Dr. Mariavittoria Masotina. One of these students, Dr. Cecilia Rocchi, conducted her research similarly to how I conducted mine, but instead focused on the case of Italy's first year of the pandemic. From her study, 7 overarching themes emerged which will help with the interpretation of the results of my research. They are:

1.3.1 **Theme of technical language:** leaders and experts often employ specific terminology and jargon which is typically not used colloquially or in non-professional settings.

1.3.2 **Theme of validation and shared responsibility:** leaders don't typically make major decisions during crises alone, so communicating with the population that they have consulted with field experts, advisors and committees before deciding on a course of action can increase credibility and diffuse potential negative repercussions.

1.3.3 **Theme of emergency as a temporary and urgent experience:** contextualized to my corpus, themes of urgency and impermanence emerge around the words *crise* ("crisis") and *moment* ("moment"). The word *moment*, especially, conveys the idea that the crisis had begun and was also close to conclusion.

1.3.4 **Theme of localization of action:** in Cecilia Rocchi's corpus (the corpus these themes were identified in), the theme of localization of action emerges due to the particularity of Italy's situation: to contrast the second wave of Covid-19 that spread all across Europe in autumn of 2020, Italy shifted gears and went from nation-wide regulations, to differentiated measures introduced on a regional basis.

1.3.5 **Theme of highlighting the error:** government leaders and prominent political figures do not act in a vacuum, they transpose each other's discourse and often incorporate its information into their own. This is especially true for the Opposition corpus, whose role could be said to be one of contestation.

1.3.6 **Theme of economy:** a health crisis of the magnitude of the Covid-19 pandemic won't only affect the healthcare systems, it will inevitably bleed into all aspects of a country's structure, and special attention will be directed, at some point, towards the economic state of said country.

1.3.7 **Theme of shift of focus, from economy to school and education:** as proven by the mere existence of the above mentioned 6 identifiable themes, during the Covid-19 pandemic, the focus, concern and resources shifted greatly based on what was perceived as being the priority at any given moment. One shift of focus that Dr. Rocchi highlights is that from economy to education.

1.3 Objective

This thesis presents a quantitative analysis of select speeches delivered by French president Emmanuel Macron, and Jean Luc Mélenchon, one of his main opposers, during the first two years of the Covid-19 pandemic in France (January 2020 – December 2021).

These political figures were selected to compare and contrast the wording used by a French government majority leader with that of their opposition. This task was carried out by a combination of a word analysis software and a computational linguistics analysis software which extracted nouns of significant frequency in both the majority's and the opposition's speeches. The same softwares were used a second time to compare the speeches of the first and the second national lockdown. With the aim of understanding the differences, these words were then sorted into thematic categories based on semantic similarity, where emergent patterns or recurrences were highlighted and commented.

With this thesis, I hope to contribute to the general understanding of the importance of (in)effective leadership messaging during times of hardship by analyzing the terminology employed by the majority and the opposition, as well as during the first and the second lockdown, in order to highlight any potential thematic differences between them.

2. METHODS

2.1 Timeline

This thesis offers an analysis of the content of some of the most important Covid-19 related speeches delivered by French President Emmanuel Macron, and one of his primary opposers, Jean-Luc Mélenchon between January 2020 and December 2021. Our first step consisted in creating a meaningful timeline of Covid-19 related events. What guided the event selection was a desire to highlight critical moments in which major political decisions were being made in France as a response to the pandemic, such as mobility restrictions, travel bans, lockdowns, the introduction of preventative measures (e.g., masks or health passes), etc. The final selection narrowed down to 10 relevant events that fit the criteria. These events were subsequently placed within a timeline of "*general periods*" (five) which aimed to highlight the recurring phenomena of these two years (confinements, extensions of confinement, deconfinements), with the exception of the final "general period" which was characterized by the availability or widespread access to the vaccine (see Table 1).

2.2 Informants

In order to perform the comparative analyses mentioned in section 1.1, two informants were selected and their speeches selected and compiled. The two informants had to be prominent political figures in France who gave speeches on the subject of Covid-19 throughout the years 2020 and 2021. They needed to represent (1) the government majority as well as (2) the opposition. When selecting these informants, Emmanuel Macron was the obvious choice for government majority leader, given that he had been the President of France since 2017 and had made or overseen all decisions regarding the handling of the pandemic since the first restrictions were put in place in the beginning of 2020. The selection of the informant for the opposition proved slightly more complex, as there were two French political figures who represented Macron's main opposers: president of the far-right nationalist party "Rassemblement National" Marine Le Pen, and the founder and head of the left-wing populist party "La France Insoumisme" Jean-Luc Mélenchon. For the purpose of this thesis, we exclusively focused on the content of speeches and disregarded any comments on the subject of Covid-19 made during interviews. This distinction was made in order to rule out interviewers' phrasing of questions, as they would inevitably affect the answer. For this reason, although both politicians had been distinctly vocal in expressing their dissent towards Macron's handling of the pandemic, Mélenchon was ultimately selected as informant for the opposition

since he had delivered a greater number of speeches in 2020 and 2021 than Le Pen, who had expressed her distain for Macron mostly during interviews.

2.3 Corpus

A *corpus* is a collection of natural language (text, and/or transcription of speech or signs) constructed with a specific purpose (Björkenstam, 2013). In this research, our corpus is comprised of 15 speeches (11 majority, 4 opposition). In addition to them covering the topic of Covid-19, the second requirement for their inclusion was that the timing of the speeches needed to coincide with the timing of the events in our timeline. Ten out of eleven transcripts of Macron's speeches were found on the official website of the French Presidency (https://www.elysee.fr/), with the one transcript (file name M3_13-04-20) being found on the government website (https://www.vie-publique.fr/). No original transcripts were found for the opposition's speeches, so video recordings of the speeches were uploaded to an automated transcription software called Sonix (https://sonix.ai/).

All 15 transcripts were then checked for errors or typos and manually cross-referenced with the video recordings of the speeches before being converted into plain text files (.txt), which were named by reporting the initial of the informant's surname, the corresponding event on the timeline, and the date (e.g., M1_12-03-20) (see Table 1).

TIMELINE OF EVENTS AND SPEECHES (JANUARY 2020 - DECEMBER 2021)					
YEAR	GENERAL PERIOD	EVENT	SPEECH BY MAJORITY	SPEECH BY OPPOSITION	TOT. SPEECHES
2020	Restrictions and first lockdown (01.20 - 05.20)	**1.** first case in France, first restrictions (24.01 - 12.03)	M1_12-03-20		4
		2. first confinement (16.03 - 13.04)	M2_16-03-20	ME2_16-03-20	
		3. extension of confinement (13.04 - 11.05)	M3_13-04-20		
	Re-opening (05.20 - 10.20)	**4.** return to "normalcy" (11.05 -28.10)	M4_14-06-20		1
	Restrictions and second lockdown (10.20 - 12.20)	**5.** second confinement (28.10 - 24.11)	M5_28-10-20	ME5_29-10-20	4
		6. extension of confinement (24.11 -15.12)	M6_24-11-20		
		7. end of the year 2020	M7_31-12-20		
2021	Restrictions and third lockdown (01.21 - 05.21)	**8.** third confinement (03.04 - 03.05)	M8_31-03-21	ME8_01-04-21	2
	Vaccination (07.21 - 12.21)	**9.** "pass sanitaire" (01.07 - 31.12)	M9_12-07-21, M9.1_09-10-21	ME9_25-07-21	3
		10. end of the year 2021	M10_31-12-21		1

Table 1, "Timeline of events and speeches (January 2020 – December 2021)": *2020-2021 timeline of important Covid-19 related events in France. The first column indicates the year, the second delineates the general period in which the specific events (column 3) took place, the fourth and fifth columns contain the file names of the speeches from our corpus of both the majority and the opposition, and the sixth column indicates the total number of speeches per general period.*

2.4 Stoplist

Once created, our corpus was uploaded and compiled on the computational linguistics software *SketchEngine* (https://www.sketchengine.eu/). This software processes texts and identifies all the instances of a word contained within them, phrase or phenomenon, and creates *word lists* or *word sketches*, compiling *concordances* therein. The software's "WordList" tool generates frequency lists of various kinds, based on parameters of frequency set by the researcher, depending on their focus (nouns, verbs, adjectives, words containing certain characteristics, word attributes etc.) The "Concordance" tool, instead, locates words, phrases, tags, documents, text types or corpus structures, and displays the contextualized results in the form of a concordance. Lastly, the "WordSketch" tool provides a summary of a word's grammatical and collocational behavior.

Cambridge Dictionary defines "lemmatization" as "the process of reducing the different forms of a word to one single form (lemma), for example, reducing 'builds', 'building', or 'built' to the lemma 'build'" (January, 2022). Using the WordList tool, I compiled a frequency list of all the lemmatized nouns that appeared in the corpus with a minimum frequency of 2 (meaning they were used at least twice across the whole corpus). The resulting frequency list had 892 items. Although this tool is algorithm based, it remains subject to error, therefore each item needed to be revisited to assess whether it was, in fact, a noun or not. Afterwards, a *"stoplist"* (Table 2) was created with all the "not-nouns" that were to be excluded from all future analyses. Once the items from our stoplist were removed, the final frequency list contained 879 items.

STOPLIST			
item	**translation**	**reason**	**observation**
arrière	back	adverb	
auto	self	prefix	
avant	before	adverb	
h	o'clock	adverb	
lorsqu'il	when it	NN	
lorsqu'ils	when they	NN	
lorsqu'on	when we	NN	
moi	me	pronoun	
parce	because	NA	parce (que)
précédent	previous	adjective	
quant	as/as to	adverb	
quoiqu'il	whatever	pronoun	
vôtre	your	possessive pronoun	

Table 2, "Stoplist": *The items are the identified errors from our frequency list compiled from all the nouns in our corpus with a minimum frequency of 2.*
Note: NN= not noun

2.5 Top 10 wordlist

The SketchEngine software provides different parameters (under the section called "viewing options") for their frequency lists. Those that fit the research objectives best were: the *absolute frequency* (how many times the item was found in the corpus), the *document frequency* (how many different documents contain the item), *relative DOCF* (the percentage of documents that contain the item), and the *ARF* (Average Reduced Frequency, which is a modified frequency whose calculation prevents the results from being excessively influenced by a high concentration of a token in only one or more small parts of the corpus). Remembering to remove the lemmas from our stoplist through the advanced parameters setting on SketchEngine, a second frequency list was created following the same criteria as before (find all nouns with a minimum frequency of 2). The items were then arranged from highest to lowest value of ARF. The first ten items from this list constitute the *"top 10 list"*, which is an itemized list of the ten most frequently found nouns in our corpus (section 3, Table 5).

The lemmas from this *top 10 list* were then defined and disambiguated (see appendix). As stated in Merriam-Webster dictionary, to "*disambiguate*" means to establish a single semantic or grammatical interpretation of a word ("Disambiguate," 2022). To define and disambiguate the ten most frequently found nouns in our corpus, SketchEngine's WordSketch tool (see 2.4) shows each instance in which the word appears and categorizes them based on the specific usage of said word (e.g., as an object of a specific verb, as the subject to a specific verb etc.). Each instance identified by the software was then manually examined, the contextualized meaning of the noun was so extrapolated and reported on, accompanied by an example which was found to be most representative of the specific shade or meaning which was being described.

2.6 Statistical analysis: log likelihood test

The statistical analyses was comprised of log-likelihood tests (G2) aimed at identifying any statistically significant differences in the frequency of the nouns used in the subcorpora to be compared: (1) Majority vs. Opposition and (2) First vs. Second lockdown.

The first step consisted in creating the subcorpora: SketchEngine allows for this to be done by clicking on the function "manage corpus", then "subcorpora" and selecting the files of the desired speeches to incorporate in the subcorpus before giving it a meaningful name (e.g., "Majority subcorpus"). This step was repeated four times in order to create four different subcorpora: twice to create the "*Majority subcorpus*" and "*Opposition subcorpus*" respectively for the first analysis ("Majority vs. Opposition"), and twice to create the "*First lockdown*

subcorpus" and "*Second lockdown subcorpus*" for the "First vs. Second lockdown" analysis (Table 3).

FILE NAMES INCLUDED IN EACH SUBCORPUS			
MAJORITY vs. OPPOSITION		FIRST vs. SECOND LOCKDOWN	
Majority Sub.	Opposition Sub.	1st Lock. Sub.	2nd Lock. Sub.
M1_12-03-20	ME2_16-03-20	M1_12-03-20	M5_28-10-20
M2_16-03-20	ME5_29-10-20	M2_16-03-20	M6_24-11-20
M3_13-04-20	ME8_01-04-21	M3_13-04-20	M7_31-12-20
M4_14-06-20	ME9_25-07-21	ME2_16-03-20	ME5_29-10-20
M5_28-10-20			
M6_24-11-20			
M7_31-03-21			
M8_31-03-21			
M9_12-07-21			
M9.1_09-10-21			
M10_31-12-21			

Table 3, "File names included in each subcorpus": *There are four columns indicating the four subcorpus that were created (two for the "Majority vs. Opposition" analysis and two for the "First vs. Second lockdown" analysis). The columns contain the names of the files belonging to each subcorpus.*

Once the necessary subcorpus were established, respective frequency lists were created with the same criteria as those detailed in section 2.5, as well as the addition of each selected subcorpus (this step had not been necessary before this point, as the frequency lists were being made on the basis of the entirety of the corpus). The resulting frequency lists were then given meaningful names (e.g., French_Majority_FreqList.csv), exported as .csv files.

In order to calculate the log-likelihood, the use of "RStudio" (https://www.rstudio.com/) was required. The version used in this thesis is RStudio 2022.07.1. The script used for running the analyses on RStudio is based on the "CorpusComparisonScript" by Ryan Boyd (https://www.ryanboyd.io/software/meh/scripts/) and was adapted by Mariavittoria Masotina to run on the output of SketchEngine.

The log-likelihood (G2) is calculated by considering the observed frequencies and the expected frequencies of the token (lemma) present in the two corpora. The test returns the G2 value, which is associated with a significance level (G2 $\geq$ 3.84, p < .05; G2 $\geq$ 6.63, p <.01; G2 $\geq$ 10.83, p <.001; G2 $\geq$ 15.13, p <.0001). The null hypothesis (H0) is that the lemma is equally frequent in the two subcorpora.

As effect size index, the Bayes Factor (BIC) was used. The interpretation is the following:

- BIC < 2 = not worth more than a bare mention. Lemmas associated with this BIC were not included in our results.
- BIC > 2 = positive evidence against H0
- BIC > 6 = strong evidence against H0
- BIC > 10 = very strong evidence against H0

2.7 Overarching themes, adapted to my corpus of data

As detailed in section 1.3, the overarching themes identified by Dr. Rocchi in her thesis were used to interpret my results. However, upon consulting with my supervisors Prof. Spagnolli and Dr. Masotina, we determined that not every thematic category identified by my colleague could be found in this corpus of data. To make these categories more representative of the corpus in question, we decided to remove the "theme of localization of action" category and adapt some of the others. The motivation for removing that particular category, lies in the fact that it was specific to Italy's handling of the pandemic: to contrast the second wave of Covid-19 that spread all across Europe in autumn 2020, Italy introduced differentiated measures on a regional basis. This approach was not adopted by France, making this theme irrelevant to our corpus. Moreover, the introduction of a category for the nouns that have no clear-cut affiliation with a theme, as well as the addition of "subthemes" to some of the categories, were necessary expansions which made the themes more coherent with this corpus. The decision to color-code the thematic categories was made along with my supervisors, and the reason behind this decision was to support the discussion of my results with a visual representation of them. Table 4 illustrates the modifications made to the original thematic categories, delineated through color-coding.

OVERARCHING THEMES IDENTIFIED BY DOTT.SSA ROCCHI, ADAPTED TO MORE ACCURATELY REFLECT THE THEMES OF MY CORPUS OF DATA		
	ORIGINAL THEME	ADAPTED THEME	SUBTHEMES
1	Theme of technical language	Theme of technical and medical language	
2	Theme of validation and shared responsibility	Theme of validation and shared responsibility	
3	Theme of emergency as a temporary and urgent experience	Theme of emergency as a temporary and urgent experience	1. Temporariness
			2. Urgency
			3. Strategy
4	Theme of localization of action	REMOVED	
5	Theme of highlighting the error	Theme of highlighting the error and the perpetrator	
6	Theme of the economy	Theme of the economy	
7	Theme of shift of focus, from the economy to school and education	Theme of school	
		No theme	

Table 4, "Overarching themes identified by Dott.ssa Rocchi, adapted to more accurately reflect the themes of my corpus of data": *the first column shows the original thematic categories identified by Dott.ssa Rocchi, the second shows how they were adapted to better represent this corpus, and the third column shows the addition of three "subthemes" to the Theme of emergency as a temporary and urgent experience*

3. RESULTS

3.1 Top 10: defined and disambiguated

Table 5 contains the ten most frequently found lemmas in our corpus, which were identified by creating a frequency list of the nouns that appeared at least twice (min. frequency of 2) across our whole corpus and arranging it from largest to smallest value of ARF (see section 2.5 in methods). In appendix, all ten lemmas were defined and disambiguated and each specific shade or meaning was accompanied by an example which was found to be most representative of it. I translated the examples from French to English and only include, for clarity, the translated version.

TOP TEN WORDS LIST (WHOLE CORPUS)						
Rank	Lemma	Translation	Frequency	DOCF	Relative DOCF	ARF
1	virus	virus	99	14	93,33333	56
2	jour	day	98	15	100	56
3	france	France	93	14	93,33333	51
4	vie	life	73	14	93,33333	45
5	pays	country	81	13	86,66667	43
6	crise	crisis	64	15	100	36
7	semaine	week	64	9	60	35
8	épidémie	epidemic	60	13	86,66667	34
9	moment	moment	53	14	93,33333	31
10	soir	evening	53	12	80	30

Table 5, "top 10 wordlist (whole corpus)": *The items (lemmas) in this table are arranged from largest to smallest value of ARF (Average Reduced Frequency) and they represent the most frequently found nouns in our corpus.*

3.2 Results of the "Majority vs. Opposition" statistical analysis

This section contains the results of the statistical analysis of "Majority vs. Opposition", which was performed on the software RStudio (the details of which can be found in section 2.6). Table 6 contains the lemmas which are significantly more frequent in the speeches by the opposition than in those by the majority. There is no corresponding table for the lemmas significantly more frequent in the majority subcorpus, as there are none.

LEMMAS SIGNIFICANTLY MORE FREQUENT IN THE OPPOSITION SUBCORPUS												
	MAJORITY				OPPOSITION				TEST RESULTS			
Lemma	Freq	Doc.Freq	ARF	Exp.Freq	Freq	Doc.Freq	ARF	Exp.Freq	G2	p	BIC	effect
temps (*time*)	31	10	16	37,96	10	4	5	3,04	11,26	< .001	2,49	positive
fois (*time*)	30	11	19	37,96	11	4	6	3,04	14,18	< .001	5,41	positive
chose (*thing*)	17	7	10	24,07	9	2	4	1,93	15,91	< .0001	7,14	strong
conseil (*advice/council*)	9	6	5	16,67	9	3	4	1,33	23,27	< .0001	14,5	very strong
pass (*pass*)	9	2	2	17,59	10	1	2	1,41	27,14	< .0001	18,37	very strong
liberté (*freedom*)	13	8	8	28,7	18	3	7	2,3	53,51	< .0001	44,74	very strong
loi (*law*)	5	4	2	11,11	7	4	5	0,89	20,9	< .0001	12,13	very strong
vote (*vote*)	3	2	2	7,41	5	2	2	0,59	15,9	< .0001	7,13	strong
société (*society*)	5	5	3	12,96	9	4	5	1,04	29,36	< .0001	20,59	very strong
classe (*class*)	3	2	2	8,33	6	2	2	0,67	20,23	< .0001	11,46	very strong
président (*president*)	6	3	2	18,52	14	4	9	1,48	49,35	< .0001	40,58	very strong
amendement (*amendment*)	0	0	0	2,78	3	3	2	0,22	15,61	< .0001	6,84	strong
dix (*ten*)	0	0	0	2,78	3	2	2	0,22	15,61	< .0001	6,84	strong
groupe (*group*)	0	0	0	2,78	3	2	2	0,22	15,61	< .0001	6,84	strong
onze (*eleven*)	0	0	0	2,78	3	3	2	0,22	15,61	< .0001	6,84	strong
salaire (*salery*)	0	0	0	2,78	3	2	2	0,22	15,61	< .0001	6,84	strong
théâtre (*theater*)	0	0	0	2,78	3	2	2	0,22	15,61	< .0001	6,84	strong
amende (*fine*)	0	0	0	3,7	4	3	2	0,3	20,82	< .0001	12,05	very strong
commission (*commission*)	0	0	0	3,7	4	2	2	0,3	20,82	< .0001	12,05	very strong
discipline (*discipline*)	0	0	0	3,7	4	2	2	0,3	20,82	< .0001	12,05	very strong
mot (*word*)	0	0	0	3,7	4	3	2	0,3	20,82	< .0001	12,05	very strong
assemblée (*assembly*)	0	0	0	4,63	5	3	3	0,37	26,02	< .0001	17,25	very strong
monsieur (*mister*)	0	0	0	4,63	5	2	2	0,37	26,02	< .0001	17,25	very strong
défense (*defense*)	0	0	0	5,56	6	3	3	0,44	31,22	< .0001	22,45	very strong
gens (*people*)	0	0	0	5,56	6	3	2	0,44	31,22	< .0001	22,45	very strong

Table 6, "lemmas significantly more frequent in the opposition subcorpus": *the first column contains the nouns significantly more frequent in the opposition subcorpus than in that of the majority. The subsequent columns show the values of frequency (Freq), document frequency (Doc.Freq), Average Reduced Frequency (ARF), and expected frequency (Exp.Freq), relatively to the first and the second subcorpus. The last four columns show the test results which are to be understood using the Bayes Factor (BIC) as effect size index (see section 2.6, paragraph 6 for detailed interpretation of BIC values)*

3.3 Results of the "First vs. Second lockdown" statistical analysis

This section details the results of the "First Lockdown vs. Second Lockdown" comparison. Table 7 contains the lemmas significantly more frequent during the first lockdown, while Table 8 contains those which were more frequent during the second lockdown.

LEMMAS SIGNIFICANTLY MORE FREQUENT DURING THE FIRST LOCKDOWN												
	FIRST LOCKDOWN				SECOND LOCKDOWN				TEST RESULTS			
Lemma	Freq	Doc.Freq	ARF	Exp.Freq	Freq	Doc.Freq	ARF	Exp.Freq	G2	p	BIC	Effect
mai (*may*)	12	1	2	6,37	0	0	0	5,63	15,19	< .0001	7,23	strong
monde (*world*)	11	4	6	5,84	0	0	0	5,16	13,92	< .001	5,97	positive
mobilisation (*mobilisation*)	10	3	5	5,31	0	0	0	4,69	12,66	< .001	4,7	positive
priorité (*priority*)	10	3	5	5,31	0	0	0	4,69	12,66	< .001	4,7	positive
distance (*distance*)	8	3	4	4,25	0	0	0	3,75	10,13	< .01	2,17	positive
etat (*state*)	8	4	6	4,25	0	0	0	3,75	10,13	< .01	2,17	positive
guerre (*war*)	8	2	2	4,25	0	0	0	3,75	10,13	< .01	2,17	positive
panique (*panic*)	8	3	5	4,25	0	0	0	3,75	10,13	< .01	2,17	positive
protection (*protection*)	8	3	4	4,25	0	0	0	3,75	10,13	< .01	2,17	positive

Table 7, "lemmas significantly more frequent during the first lockdown": *first column contains the nouns significantly more frequent during the first lockdown.. The subsequent columns show the values of frequency (Freq), document frequency (Doc.Freq), Average Reduced Frequency (ARF), and expected frequency (Exp.Freq), relatively to the first and the second subcorpus. The last four columns show the test results, which are to be understood using the Bayes Factor as effect size index (see section 2.6 for detailed interpretation of BIC values)*

LEMMAS SIGNIFICANTLY MORE FREQUENT DURING THE SECOND LOCKDOWN												
	FIRST LOCKDOWN				SECOND LOCKDOWN				TEST RESULTS			
Lemma	Freq	Doc.Freq	ARF	Exp.Freq	Freq	Doc.Freq	ARF	Exp.Freq	G2	p	BIC	Effect
commerce (*commerce*)	0	0	0	3,72	7	2	3	3,28	10,6	< .01	2,65	positive
décembre (*december*)	0	0	0	3,72	7	3	3	3,28	10,6	< .01	2,65	positive
emploi (*employment*)	0	0	0	3,72	7	3	3	3,28	10,6	< .01	2,65	positive
étape (*stage*)	0	0	0	3,72	7	2	3	3,28	10,6	< .01	2,65	positive
forme (*shape*)	0	0	0	3,72	7	2	4	3,28	10,6	< .01	2,65	positive
janvier (*january*)	0	0	0	3,72	7	1	2	3,28	10,6	< .01	2,65	positive
liberté (*freedom*)	0	0	0	3,72	7	2	2	3,28	10,6	< .01	2,65	positive
lit (*bed*)	0	0	0	3,72	7	3	3	3,28	10,6	< .01	2,65	positive
rassemblement (*gathering*)	0	0	0	3,72	7	3	3	3,28	10,6	< .01	2,65	positive
rôle (*role*)	0	0	0	3,72	7	2	3	3,28	10,6	< .01	2,65	positive
compte (*account*)	0	0	0	4,25	8	3	5	3,75	12,12	< .001	4,16	positive
génération (*generation*)	0	0	0	4,25	8	4	5	3,75	12,12	< .001	4,16	positive
terme (*term*)	0	0	0	4,25	8	2	3	3,75	12,12	< .001	4,16	positive
lieu (*place*)	0	0	0	4,78	9	3	6	4,22	13,63	< .001	5,67	positive
test (*test*)	0	0	0	5,31	10	2	4	4,69	15,14	< .0001	7,19	strong
contamination (*contamination*)	0	0	0	6,37	12	3	7	5,63	18,17	< .0001	10,22	very strong
printemps (*spring*)	0	0	0	6,9	13	3	6	6,1	19,69	< .0001	11,73	very strong
stratégie (*strategy*)	0	0	0	6,9	13	3	7	6,1	19,69	< .0001	11,73	very strong
an (*year*)	0	0	0	7,43	14	4	7	6,57	21,2	< .0001	13,25	very strong
vague (*wave*)	0	0	0	7,43	14	3	7	6,57	21,2	< .0001	13,25	very strong

Table 8, "lemmas significantly more frequent during the second lockdown": *frequency list of the nouns significantly more frequent in the second lockdown subcorpus than in the first.*

3.4 Categorization of nouns by theme

In section 2.7 I explained how the iteration of themes that best represent this corpus was created and on what basis the color-coding method was determined. In this section, I will include a table (Table 9) with the lemmas from Table 5, Table 6, Table 7, and Table 8, sorted into categories, along with meaningful instances which support and contextualize the decision for their thematic affiliation.

NOUNS FROM TABLES 5,6,7,8 ORGANIZED BY THEMATIC AFFILIATION				
	FREQ. LIST	**LEMMA**	**EXAMPLE**	**SUBTHEMES**
1. THEME OF TECHNICAL AND MEDICAL LANGUAGE	top 10	**virus *(virus)***	*"For several weeks, our country has been facing the spread of a **virus**, Covid-19, which has affected several thousand of our compatriots", M1*	
	top 10	**épidémie *(epidemic)***	*"Thanks to the health pass and the strategy implemented since last July, we have managed to control the **epidemic**", M9.1*	
	sig. freq. opp.	**pass *(pass)***	*"Thanks to the health **pass** and the strategy implemented since last July, we have managed to control the epidemic", M9.1*	
	sig. freq. opp.	**loi *(law/bill)***	*"We would not ask you to adopt our eleven **laws**, our five plans to address the health emergency, or the findings of our commission of inquiry. We would not ask you to make amends for laughing in our faces when we warned you about a possible second wave, nor for the arrogant answers you gave us throughout the discussion of the Social Security budget **bill**.", ME5*	
	sig. freq. opp.	**amendement (amendement)**	*"You present us with a plan, a take-it-or-leave-it plan without a word, without a possible **amendment** or a line. You ask us for a vote of confidence? Well, we don't have confidence", ME5*	
	sig. freq. 1 lockd.	**distance *(distance)***	*"All companies must organize themselves to facilitate **remote** work", M2*	
	sig. freq. 2 lockd.	**contamination *(contamination)***	*"On December 15, if we have reached around 5,000 **contaminations** per day and around 2,500 to 3,000 people in intensive care, we will be able to reach a new milestone. Then the confinement can be lifted", M6*	
	sig. freq. 2 lockd.	**rassemblement (gathering)**	*"limit all activities that multiply **gatherings**, lead people to gather without protection in closed places and progressively allow to reopen activities where one can protect*	

			oneself and where distances and barrier gestures are possible", M6	
	sig. freq. 2 lockd.	**test** *(test)*	*"Today, according to the first data which will soon be refined by the so-called serological **tests**, a very small minority of French people have contracted Covid-19.", M3*	
	sig. freq. 2 lockd.	**forme** *(shape)*	*"On the other hand, the virus develops and develops severe **forms** in younger people.", M5*	
	sig. freq. 2 lockd.	**lit** *(bed)*	*"The number of intensive care **beds** has already been increased to 7000.", M8*	
2. THEME OF VALIDATION AND SHARED RESPONSIBILITY	sig. freq. opp.	**défense** *(defense)*	*"So, Mr. Castex, are you the factor? The Council of **Defense**, said yesterday the spokesman for the government.", ME5*	
	sig. freq. opp.	**assemblée** *(assembly)*	*"But who is the author of this plan? Not the National **Assembly**, not its committee. Not the government either.", ME5*	
	sig. freq. opp.	**conseil** *(council)*	*"It is up to the government to plan actions after consulting its Scientific and Medical **Councils**.", ME5 "This is the fight that I will lead on your behalf from the European **Council** in July and in the next two years.", M4* (also Economic, Defense, and Ministers Councils)	
	sig. freq. opp.	**commission** *(commission)*	*"But who is the author of this plan? Not the National Assembly, not its **commission**.", ME5*	
	sig. freq. opp.	**vote** *(vote)*	*"From tomorrow, a debate followed by a **vote** will be held in Parliament. Tomorrow, the Government will detail all these measures, and we must be proud of this transparency and of the fact that these difficult decisions are made in a democratic framework, where all the oppositions can express themselves.", M5*	
	sig. freq. 1 lockd.	**etat** *(state)*	*"I trust the mayors and the civic-mindedness of each of you. I also know that the town halls and the **State** services have organized things well", M1*	
3. THEME OF EMERGENCY AS A TEMPORARY AND URGENT EXPERIENCE	top 10	**jour** *(day)*	*"My dear compatriots, we will have better **days** and we will find the happy **days**" M3*	temporariness
	top 10	**vie** *(life)*	*"They – the frontline workers- all accepted to take time out of their personal and family **lives** for our health.", M1*	urgency
	top 10	**crise** *(crisis)*	*"People of France, Thursday night, I addressed you to talk about the health **crisis** our country is going through. Until then, the*	urgency

			epidemic may have been a distant thought for some…", M2	
	top 10	**semaine** *(week)*	*"During the next 4 **weeks**, the rules provided by the government must continue to be respected", M3*	temporariness
	top 10	**moment** *(moment)*	*: "But compared to this same **moment** last year when the constraints were much stronger, we have the weapon of the vaccine, and our collective experience. And therefore, real reasons for hope", M10*	temporariness
	top 10	**soir** *(evening)*	*"…the rules which are currently in force in the 19 'reinforced vigilance' departments, these rules will be extended to all the metropolitan territory starting from this Saturday **evening** and for 4 weeks", M8*	temporariness
	sig. freq. opp.	**société** *(society)*	*"By fighting against the virus, by fighting against its economic and social consequences, by rebuilding a stronger, more fraternal and sustainable **society**, we will build the France of 2030 ", M7*	strategy
	sig. freq. opp.	**discipline** *(discipline)*	*I might as well tell you that we will respect the health **disciplines** that you will have adopted here by the vote of your majority. ME5*	strategy
	sig. freq. opp.	**temps** *(time)*	*"We must continue to gain **time**, and to do so, I am going to ask you to continue to make sacrifices and rather to do more, but for our collective interest", M1*	temporariness
	sig. freq. 1 lockd.	**mai** *(may)*	*"From **May** 11, we will gradually reopen the nurseries, schools, colleges and high schools", M3*	temporariness
	sig. freq. 1 lockd.	**guerre** *(war)*	*"We are at **war**, in a health war, certainly: we are not fighting against an army, nor against another Nation. But the enemy is there, invisible, elusive, advancing", M2*	urgency
	sig. freq. 1 lockd.	**panique** *(panic)*	*"There is no room for complacency or **panic**. We will undoubtedly have to take control measures, close borders …", M1*	urgency
	sig. freq. 1 lockd.	**protection** *(protection)*	*"We will then deploy all the necessary means to **protect** the population", M3*	strategy
	sig. freq. 1 lockd.	**priorité** *(priority)*	*"we are only at the beginning of this epidemic … Faced with this, the top **priority** for our Nation will be our health.", M1*	strategy
	sig. freq. 1 lockd.	**mobilisation** *(mobilisation)*	*"The general **mobilization** is also that of our researchers. Numerous French and European programs and clinical trials are underway to produce large quantities of*	strategy

			rapid, effective and efficient diagnostics", M1	
	sig. freq. 2 lockd.	**printemps** *(spring)*	*"Hope is there, and hope is growing every day in this recovery that is already simmering in France more than anywhere else and that will allow us, starting in the* **spring***, to invent a stronger economy …",* M7	temporariness
	sig. freq. 2 lockd.	**stratégie** *(strategy)*	*"Only then will we be able to redeploy a renewed and completed 'Test, Alert, Protect'* **strategy***",* M5	strategy
	sig. freq. 2 lockd.	**terme** *(term)*	*"Given that the other European countries are saturated, we will not be able to call upon a foreign workforce in the short* **term***",* M5	temporariness
	sig. freq. 2 lockd.	**décembre** *(december)*	*"All these measures will be applied as soon as possible. They will be implemented during the night from Thursday to Friday, and will be applied at least until* **December** *1st.",* M5	temporariness
	sig. freq. 2 lockd.	**étape** *(stage)*	*"A new* **stage** *will open on Saturday morning, November 28. First, there is what will not change: adaptive containment and therefore the certification system will remain in place because that is what has allowed us to achieve these good results",* M6	temporariness
	sig. freq. 2 lockd.	**janvier** *(january)*	*"The next step will be for* **January** *20. Indeed, on this date, we will have enough time to look back at the end of the year",* M6	temporariness
	sig. freq. 2 lockd.	**liberté** *(freedom)*	*"How long, at what time we must return and how much we will be fined if we disobey him? What is left of our individual* **freedom***? Foundational to human identity. What happened to our collective freedoms?",* ME5	urgency
	sig. freq. 2 lockd.	**rôle** *(role)*	*"This new confinement will only succeed if everyone is mobilized and everyone has a* **role** *to play.",* M5	strategy
	sig. freq. 2 lockd.	**lieu** *(place)*	*"To do this, we must continue for several weeks to avoid many activities in enclosed* **places** *that accelerate the spread of the virus, and limit gatherings to a minimum.",* M6	strategy
	sig. freq. 2 lockd.	**vague** *(wave)*	*"Unlike the first* **wave***, all regions are now at the alert threshold.",* M5	temporariness
4. THEME OF HIGHLIGHTING THE ERROR	sig. freq. opp.	**président** *(president)*	*"'You will only go out to work' said the* **president***. We will live from fortnight to fortnight. While we wait for the decisions of the Defense Council and* **President** *Macron,*	

AND THE PERPETRATOR			*who will tell us at what time we can go out?", ME5*
	sig. freq. opp.	**monsieur** *(mister)*	*"Where did the plan he presented last night come from? You're repeating it today in your speech, **mister** Prime Minister. But who is the author of this plan?", ME5*
	sig. freq. opp.	**amende** *(fine)*	*"How long, at what time we must return and how much we will be **fined** if we disobey him?", ME5*
	sig. freq. opp.	**mot** *(word)*	*"What are the **words** that are spoken worth? That's what we'll see in the coming hours.", ME2*
5. THEME OF ECONOMY	sig. freq. opp.	**salaire** *(salery)*	*"Unvaccinated caregiver? Suspended. He must learn to live without a **salery**.", ME9*
	sig. freq. 2 lockd.	**commerce** *(commerce/business)*	*"The **businesses** that were defined in the spring as non-essential, establishments receiving the public, including bars and restaurants, will be closed.", M5*
	sig. freq. 2 lockd.	**emploi** *(employment)*	*"Thirdly, it is a question of protecting the poorest people who, because they live in more cramped quarters, because they have precarious **employment**, are the most affected by the virus in terms of health, but they are also the most affected by the economic and social consequences of the crisis.", M5*
6. THEME OF SCHOOL	sig. freq. opp.	**classe** *(class)*	*"You began by saying that there was no danger to the children, then that the danger existed to the point that one sick person per **class** justified, confirmed the class, then that it could be three. It doesn't take a rocket scientist to figure out that if three were sick, all the others would soon be sick, along with their families.", ME8*
7. NO THEME	top 10	**france** *(france)*	*"A united **France** is our best asset in the troubled times we are going through. We will all stand together. Long live the Republic!" M1 (no identifiable theme)*
	top 10	**pays** *(country)*	*"Our **country** will invest even more heavily in research", M3 (no identifiable theme)*
	sig. freq. opp.	**onze** *(eleven)*	*"We would not ask you to adopt our **eleven** bills, our five plans to address the health emergency, or ... to make amends for laughing in our faces when we warned you about a possible second wave.", ME5 (number)*

	sig. freq. opp.	**gens** *(people)*	*"And it won't prevent **people** from infected, it won't prevent **people** from getting sick.", ME2 (no identifiable theme)*	
	sig. freq. opp.	**chose** *(thing)*	*"And it won't prevent people from getting infected, it won't prevent people from getting sick. We have to call **things** as they are.", ME2 (abstract figure of speech)*	
	sig. freq. opp.	**dix** *(ten)*	*"Since then, there have been **ten** plays on this theme, ten films made since '61, eleven operas since 1762.", ME9 (number)*	
	sig. freq. opp.	**groupe** *(group)*	*"You have not provided for any hiring of teacher assistants to help teachers organize academic remediation. And the division of the workforce into sub**groups**.", ME5 (no identifiable theme)*	
	sig. freq. opp.	**théâtre** *(theater/play)*	*" Antigone returns! Do you know the story about her? ... Antigone decides, at the risk of her own life, to bury her brother, condemned for his religion, the rebellion ... Here is the universal and permanent lesson. Respect for the right of each person is our first duty. Since then, there have been ten **plays** on this theme, ten films made since 1961, eleven operas since 1762", ME9 (irrelevant noun)*	
	sig. freq. opp.	**fois** *(time)*	*"That's why once again, I'm calling on the 5 million unvaccinated. Make this simple gesture. For you. For your fellow citizens. For our country.", M10 (figure of speech)*	
	sig. freq. 1 lockd.	**monde** *(world)*	*"we also saw people gathering in the parks, crowded markets, restaurants, bars that did not respect the closing instructions.", M2 (figure of speech)*	
	sig. freq. 1 lockd.	**etat** *(state)*	***state** of France, state bodies ("I trust the mayors and the civic-mindedness of each of you. I also know that the town halls and the State services have organized things well", M1)*	
	sig. freq. 2 lockd.	**an** *(year)*	*Tonight, I ask all elderly people over 70 **years** old, ... , to stay at home as much as possible.", M1 (too many unimportant meanings)*	
	sig. freq. 2 lockd.	**compte** *(account)*	*"taking into **account** that other European countries are saturated, we will not be able to call on a foreign workforce in the short term.", M5 (figure of speech)*	

	sig. freq. 2 lockd.	**génération** *(generation)*	*" It therefore affects in serious forms, all* **generations** *", M5 (<u>no identifiable theme</u>)*	

Table 9: *Lemmas from Table 5, Table 6, Table 7, and Table 8, sorted into thematic categories, along with meaningful instances which support and contextualize the sorting*

4. Discussion

In this section, color-coded versions of Tables 5 through 8 (now Table 10, Table 11, Table 12, and Table 13) will provide a visual representation of the recurrence of themes in certain subcorpus, which will support the discussion of my results.

4.1 Themes in top 10 (Table 10)

TOP TEN WORDS LIST (WHOLE CORPUS)						
Rank	Lemma	Translation	Frequency	DOCF	Relative DOCF	ARF
1	virus	virus	99	14	93,33333	56
2	jour	day	98	15	100	56
3	france	France	93	14	93,33333	51
4	vie	life	73	14	93,33333	45
5	pays	country	81	13	86,66667	43
6	crise	crisis	64	15	100	36
7	semaine	week	64	9	60	35
8	épidémie	epidemic	60	13	86,66667	34
9	moment	moment	53	14	93,33333	31
10	soir	evening	53	12	80	30

Table 10: *Table 5, color-coded.*
Purple: "theme of technical and medical language"; Red: "no theme"; Green: "theme of emergency as a temporary and urgent experience" (with subthemes "temporariness": light green, and "urgency": dark green)

Six out of the ten most frequently used nouns in our whole corpus fall under the category "theme of emergency as a temporary and urgent experience" (green). It is interesting to note that, among the nouns belonging to this thematic category, the "strategy" subtheme is absent, with four of these words being associated with "temporariness" and the other two with "urgency". This suggests that overall, throughout the pandemic, reassuring the population about the temporary nature of this health-crisis, took priority over informing them of the strategy that was being employed to actually contrast it.

4.2 Themes in the opposition subcorpus (Table 11)

Although in this subcorpus the "theme of highlighting the error and the perpetrator" category (yellow) includes only about 17% of the lemmas from this list, Table 9 demonstrates that, in fact, the opposition subcorpus is the only subcorpus that is represented by this theme. The informants selected in this analysis do not act in an abstract context, they transpose each other's discourse and incorporate its information into their own: this is especially true for the opposition, whose role could be said to be one of contestation to the majority, which explains

why they would be more inclined to highlight the errors committed than would be the majority itself.

LEMMAS SIGNIFICANTLY MORE FREQUENT IN THE OPPOSITION SUBCORPUS												
	MAJORITY				OPPOSITION				TEST RESULTS			
Lemma	Freq	Doc.Freq	ARF	Exp.Freq	Freq	Doc.Freq	ARF	Exp.Freq	G2	p	BIC	effect
liberté (*freedom*)	13	8	8	28,70	18	3	7	2,30	53,51	< .0001	44,74	very strong
président (*president*)	6	3	2	18,52	14	4	9	1,48	49,35	< .0001	40,58	very strong
défense (*defense*)	0	0	0	5,56	6	3	3	0,44	31,22	< .0001	22,45	very strong
gens (*people*)	0	0	0	5,56	6	3	2	0,44	31,22	< .0001	22,45	very strong
société (*society*)	5	5	3	12,96	9	4	5	1,04	29,36	< .0001	20,59	very strong
pass (*pass*)	9	2	2	17,59	10	1	2	1,41	27,14	< .0001	18,37	very strong
assemblée (*assembly*)	0	0	0	4,63	5	3	3	0,37	26,02	< .0001	17,25	very strong
monsieur (*mister*)	0	0	0	4,63	5	2	2	0,37	26,02	< .0001	17,25	very strong
conseil (*advice/council*)	9	6	5	16,67	9	3	4	1,33	23,27	< .0001	14,5	very strong
loi (*law*)	5	4	2	11,11	7	4	5	0,89	20,9	< .0001	12,13	very strong
amende (*fine*)	0	0	0	3,70	4	3	2	0,30	20,82	< .0001	12,05	very strong
commission (*commission*)	0	0	0	3,70	4	2	2	0,30	20,82	< .0001	12,05	very strong
discipline (*discipline*)	0	0	0	3,70	4	2	2	0,30	20,82	< .0001	12,05	very strong
mot (*word*)	0	0	0	3,70	4	3	2	0,30	20,82	< .0001	12,05	very strong
classe (*class*)	3	2	2	8,33	6	2	2	0,67	20,23	< .0001	11,46	very strong
chose (*thing*)	17	7	10	24,07	9	2	4	1,93	15,91	< .0001	7,14	strong
vote (*vote*)	3	2	2	7,41	5	2	2	0,59	15,9	< .0001	7,13	strong
amendement (*amendment*)	0	0	0	2,78	3	3	2	0,22	15,61	< .0001	6,84	strong
dix (*ten*)	0	0	0	2,78	3	2	2	0,22	15,61	< .0001	6,84	strong
groupe (*group*)	0	0	0	2,78	3	2	2	0,22	15,61	< .0001	6,84	strong
onze (*eleven*)	0	0	0	2,78	3	3	2	0,22	15,61	< .0001	6,84	strong
salaire (*salery*)	0	0	0	2,78	3	2	2	0,22	15,61	< .0001	6,84	strong
théâtre (*theater*)	0	0	0	2,78	3	2	2	0,22	15,61	< .0001	6,84	strong
fois (*time*)	30	11	19	37,96	11	4	6	3,04	14,18	< .001	5,41	positive
temps (*time*)	31	10	16	37,96	10	4	5	3,04	11,26	< .001	2,49	positive

Table 11: *Table 6, color-coded*
Purple: "theme of technical and medical language"; Red: "no theme"; Green: "theme of emergency as a temporary and urgent experience" (with subthemes "temporariness": light green, "strategy": vibrant green, and "urgency": dark green); Blue: "theme of validation and shared responsibility"; Pink: "theme of school"; Yellow: "theme of highlighting the error and the perpetrator"; Orange: "theme of economy"

Another thing that emerges from Table 11 is the consistent presence of themeless nouns (red): this can be explained by assuming that, while a government majority leader whose role is to represent the population, would employ a formal and curated register when delivering a speech, the opposition would tend to use more colloquial terminology (e.g., "*And it won't prevent people from getting infected, it won't prevent people from getting sick. We must call <u>things</u> as they are.*") or even employ irony and rhetorical figures to gain mass approval, like in the following example: "*Antigone returns! Do you know the story about her? ... Antigone decides, at the risk of her own life, to bury her brother, condemned for his religion, the rebellion ... Here is the universal and permanent lesson. Respect for the right of each person is our first duty. Since then, there have been ten plays on this theme, ten films made since 1961, eleven operas since 1762*".

4.3 Themes during the first lockdown and during the second lockdown (Tables 12 and 13)

LEMMAS SIGNIFICANTLY MORE FREQUENT DURING THE FIRST LOCKDOWN												
	FIRST LOCKDOWN				SECOND LOCKDOWN				TEST RESULTS			
Lemma	Freq	Doc.Freq	ARF	Exp.Freq	Freq	Doc.Freq	ARF	Exp.Freq	G2	p	BIC	Effect
mai (*may*)	12	1	2	6,37	0	0	0	5,63	15,19	< .0001	7,23	strong
monde (*world*)	11	4	6	5,84	0	0	0	5,16	13,92	< .001	5,97	positive
mobilisation (*mobilisation*)	10	3	5	5,31	0	0	0	4,69	12,66	< .001	4,7	positive
priorité (*priority*)	10	3	5	5,31	0	0	0	4,69	12,66	< .001	4,7	positive
distance (*distance*)	8	3	4	4,25	0	0	0	3,75	10,13	< .01	2,17	positive
etat (*state*)	8	4	6	4,25	0	0	0	3,75	10,13	< .01	2,17	positive
guerre (*war*)	8	2	2	4,25	0	0	0	3,75	10,13	< .01	2,17	positive
panique (*panic*)	8	3	5	4,25	0	0	0	3,75	10,13	< .01	2,17	positive
protection (*protection*)	8	3	4	4,25	0	0	0	3,75	10,13	< .01	2,17	positive

Table 12: *Table 7, color-coded*
Purple: *"theme of technical and medical language"; Red: "no theme"; Green: "theme of emergency as a temporary and urgent experience" (with subthemes "temporariness": light green, "urgency": dark green, and "strategy": vibrant green)*

LEMMAS SIGNIFICANTLY MORE FREQUENT DURING THE SECOND LOCKDOWN												
	FIRST LOCKDOWN				SECOND LOCKDOWN				TEST RESULTS			
Lemma	Freq	Doc.Freq	ARF	Exp.Freq	Freq	Doc.Freq	ARF	Exp.Freq	G2	p	BIC	Effect
an (*year*)	0	0	0	7,43	14	4	7	6,57	21,2	< .0001	13,25	very strong
vague (*wave*)	0	0	0	7,43	14	3	7	6,57	21,2	< .0001	13,25	very strong
printemps (*spring*)	0	0	0	6,9	13	3	6	6,1	19,69	< .0001	11,73	very strong
stratégie (*strategy*)	0	0	0	6,9	13	3	7	6,1	19,69	< .0001	11,73	very strong
contamination (*contamination*)	0	0	0	6,37	12	3	7	5,63	18,17	< .0001	10,22	very strong
test (*test*)	0	0	0	5,31	10	2	4	4,69	15,14	< .0001	7,19	strong
lieu (*place*)	0	0	0	4,78	9	3	6	4,22	13,63	< .001	5,67	positive
compte (*account*)	0	0	0	4,25	8	3	5	3,75	12,12	< .001	4,16	positive
génération (*generation*)	0	0	0	4,25	8	4	5	3,75	12,12	< .001	4,16	positive
terme (*term*)	0	0	0	4,25	8	2	3	3,75	12,12	< .001	4,16	positive
commerce (*commerce*)	0	0	0	3,72	7	2	3	3,28	10,6	< .01	2,65	positive
décembre (*december*)	0	0	0	3,72	7	3	3	3,28	10,6	< .01	2,65	positive
emploi (*employment*)	0	0	0	3,72	7	3	3	3,28	10,6	< .01	2,65	positive
étape (*stage*)	0	0	0	3,72	7	2	3	3,28	10,6	< .01	2,65	positive
forme (*shape*)	0	0	0	3,72	7	2	4	3,28	10,6	< .01	2,65	positive
janvier (*january*)	0	0	0	3,72	7	1	2	3,28	10,6	< .01	2,65	positive
liberté (*freedom*)	0	0	0	3,72	7	2	2	3,28	10,6	< .01	2,65	positive
lit (*bed*)	0	0	0	3,72	7	3	3	3,28	10,6	< .01	2,65	positive
rassemblement (*gathering*)	0	0	0	3,72	7	3	3	3,28	10,6	< .01	2,65	positive
rôle (*role*)	0	0	0	3,72	7	2	3	3,28	10,6	< .01	2,65	positive

Table 13: *Table 8, color-coded*
Purple: *"theme of technical and medical language"; Red: "no theme"; Green: "theme of emergency as a temporary and urgent experience" (with subthemes "temporariness": light green, and "strategy": vibrant green); Orange: "theme of economy"*

The majority (62,5%) of lemmas significantly more frequent in the subcorpus of the first lockdown showed thematic affiliation with the theme of emergency as a temporary and urgent experience (green). If we compare this to the contents of Table 13 (which contains the lemmas significantly more frequent during the second lockdown), we can notice that, although there is still a slight majority of green, it is counterbalanced by a much higher number of different themes. This can be explained by the fact that the second lockdown was, in fact, a *second* lockdown, meaning that there was more information regarding the virus than there was the first time, and that people had a greater understanding of what a lockdown actually entailed. For this reason, informing the population about the severity of the health crisis wasn't as necessary as it was during the first lockdown, when, instead, the subthemes of "strategy" and "urgency" were at the forefront of discourse. During the second lockdown, people needed to be reassured about the temporariness of the restrictions, and panic regarding the actual virus shifted towards more tangible aspects of life, such as the inevitable financial and unemployment crises caused by the pandemic, as illustrated by the following example: "*Thirdly, it is a question of protecting the poorest people who, because they live in more cramped quarters, because they have precarious employment, are the most affected by the virus in terms of health, but they are also the most affected by the economic and social consequences of the crisis*" (M5).

5. Conclusion

This thesis presented a quantitative analysis of the most important speeches delivered by French president Emmanuel Macron, and by one of his main opposers, Jean Luc Mélenchon, during the first two years of the Covid-19 pandemic in France (January 2020-December 2021). A statistical analysis of the terminology employed by my two informants, as well as that employed during the first and the second lockdown, allowed me to extract all the nouns of significant frequency within my subcorpus, which I was then able to categorize by thematic affiliation using Dr. Rocchi's seven categories as a starting point. Along with my supervisors, I adjusted the original categories to better reflect my corpus, which resulted in six overarching themes. These themes allowed for a discussion on the thematic differences between the first and second lockdown, as well as considerations on the overall discourse surrounding Covid-19 in France to take place.

Overall, the theme of emergency as a temporary and urgent experience was the most present one in our corpus. This conclusion is in line with what emerges from the "First vs. Second lockdown" comparison, with an important distinction being made between the predominant subthemes: during the first lockdown, the emphasis was on strategy, while during the second lockdown, political figures strived to reassure the public and remind them about the temporary nature of the pandemic.

Future research on this topic should focus on identifying overarching themes in other countries and investigating whether they are consistent or not with those found in this and Dr. Rocchi's theses.

References

Björkenstam, K. N. (2013). *What is a corpus and why are corpora important tools?* www.wordpress.com. https://nordiskateckensprak.files.wordpress.com/2014/01/knb_whatisacorpus_cph-2013_outline.pdf

COVID-19 - Situation épidémiologique | Coronavirus Covid-19. (2020, January 1). Sciensano. https://covid-19.sciensano.be/fr/covid-19-situation-epidemiologique

Darques, R. (2022, July 1). *Clustering and mapping the first COVID-19 outbreak in France - BMC Public Health*. BioMed Central. https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-022-13537-7

Desson, Z., Weller, E., McMeekin, P., & Ammi, M. (2020). An analysis of the policy responses to the COVID-19 pandemic in France, Belgium, and Canada. *Health Policy and Technology*, *9*(4), 430–446. https://doi.org/10.1016/j.hlpt.2020.09.002

disambiguate. (2022). In *Merriam-Webster*. https://www.merriam-webster.com/dictionary/disambiguate

Gallo, C. (2021, February 2). *Finding the Right Words in a Crisis*. Harvard Business Review. https://hbr.org/2020/04/finding-the-right-words-in-a-crisis

NEWS WIRES. (2021, January 30). *Covid-19: Macron defies scientific advisers and says "non" to lockdown*. France 24. https://www.france24.com/en/europe/20210130-macron-defends-decision-not-to-impose-lockdown-in-france-as-virus-cases-rise

Noël, R. (2020, March 11). *INFOGRAPHIE - Coronavirus : comment le Grand Est est devenu la région la plus touchée de France*. Ici, Par France Bleu Et France 3. https://www.francebleu.fr/infos/sante-sciences/infographie-coronavirus-comment-le-grand-est-est-devenu-la-region-la-plus-touchee-de-france-1583956926

Peillon, L. (2020, May 21). *L'origine de l'épidémie de Covid en France peut-elle remonter à l'automne 2019 ?* Liberation. https://www.liberation.fr/checknews/2020/05/21/l-origine-de-l-epidemie-de-covid-en-france-peut-elle-remonter-a-l-automne-2019_1788866/

Rocchi, C., Spagnolli, A., & Masotina, M. (n.d.). *Government's and opposition's public speeches on COVID-19 pandemic measures in 2020: a qualitative analysis* [MA Thesis]. Università degli studi di Padova.

RStudio (2022.07.1). (2022). [Script based on the "CorpusComparisonScript" by Ryan Boyd, adapted by Mariavittoria Masotina to interpret the output of SketchEngine]. https://www.rstudio.com/products/rstudio/download/

The New York. (2022, August 24). *Tracking Coronavirus in France: Latest Case Count*. The New York Times. https://www.nytimes.com/interactive/2021/world/france-covid-cases.html

Ward, A. (2020, April 15). *How France's President Emmanuel Macron bungled his coronavirus response*. Vox. https://www.vox.com/2020/4/14/21218927/coronavirus-covid-france-macron-response

<u>Appendix</u>

***3.1.1 Virus* (virus):** A virus is an infectious agent of small size and simple composition that can multiply only in living cells of animals, plants, or bacteria (Wagner, 2022). The noun was found 99 times within the corpus across 14 of the 15 total documents. In our corpus, the term *virus* emerges synonymously with the **disease Covid-19**, as demonstrated by the following example: "*For several weeks, our country has been facing the spread of a virus, <u>Covid-19</u>, which has affected several thousand of our compatriots*" (M1).

***3.1.2 Jour* (day):** In our corpus, the word *jour* (in English "day") can be found in all 15 documents, 98 times in total. This word is used to indicate a unit of time in the past, present or future. Whether it be **difficult days** ("*We are living <u>difficult days</u>. We are all feeling the fear, the anxiety for our parents, for ourselves in the face of this dreaded, invisible, unpredictable virus*", M3) or **better days** ("*My dear compatriots, we will have <u>better days</u> and we will find the happy days*" M3), **coming days** ("*We need to do everything we can in the <u>coming days</u> for the next two weeks to have as little contact with others as possible, so that we can slow down the spread of the contamination as much as possible*", ME2) or **previous days** ("*My dear compatriots, if I wanted to address you this evening, after having consulted widely over the <u>last few days</u>, it is to tell you in all transparency what awaits us in the coming weeks and months*", M3).

***3.1.3 France* (France):** *France* is the country of reference across our corpus, and the home country of both of our informants. It appears 93 times in 14 out of 15 total documents: *"A united <u>France</u> is our best asset in the troubled times we are going through. We will all stand together. Long live the Republic!*" (M1).

***3.1.4 Vie* (life):** *Vie*, in English "life", appears 73 times in all but one document from our corpus. The term is used most frequently to indicate a **personal life** ("*They – the frontline workers- all accepted to take time out of their <u>personal and family lives</u> for our health.*", M1), **daily life** (*"For our <u>daily life</u>, we must continue to apply the "barrier gestures" when we go out: keep our distance and wash our hands*", M3), **democratic life** ("*But it is important, at this moment, by following the advice of scientists as we have just done, to ensure the continuity of our <u>democratic life</u> and our institutions*", M1), or **financial life** *("For <u>financial life</u>, as far as France is concerned, no company, whatever its size, will be left to the risk of bankruptcy*", M2) that is subject to change because of the pandemic. In concordance with the verb *gagner* ("to win"), we find this noun twice in our corpus to indicate **earning a living**: "*This is why the unemployment insurance reform will be fully implemented as of October 1. Around a simple*

will: in France, one must always <u>earn a</u> better <u>living</u> by working than by staying at home, which is currently not always the case" (M9).

***3.1.5 Pays* (country)**: the word appears 81 times in 13 documents of our corpus, most often in reference to the **home country** of **France**, as illustrated by the following examples: "*<u>French citizens</u> who are currently abroad and wish to return will of course be able to join <u>their country</u>*" (M2), and *"<u>Our country</u> will invest even more heavily in research*" (M3). It is used 7 times across our corpus to indicate **European countries** ("*This epidemic, which affects all continents and hits all <u>European countries</u>, is the most serious health crisis that France has experienced in a century*", M1), twice for **non-European countries** (*"Until further notice, our borders with <u>non-European countries</u> will remain closed*", M3), and once for **neighboring countries** ("*We were able to double our resuscitation capacities in a few days, organize transfers of hundreds of patients between regions and with <u>neighboring countries</u>*", M4)

***3.1.6 Crise* (crisis):** The French word for "crisis" ("*crise*") appears in every document of our corpus, most frequently indicating, as expected, the **French** and **global health crisis** caused by Covid-19: "*People of France, Thursday night, I addressed you to talk about the <u>health crisis</u> our country is going through. Until then, the epidemic may have been a distant thought for some…*" (M2), and "*It is not division that will provide the answer to what is now a <u>global crisis</u>, but our ability to see right and early together and act together*" (M1)..

In some cases, it refers to the **socioeconomic crisis** caused by the pandemic: "*At the same time, we are facing an unprecedented <u>economic and social crisis</u>. It is the consequence of the first confinement and of what we are living*" (M6).

Once in our corpus, the opposition spoke of a **political crisis** caused by the introduction of the "*pass sanitaire*", a "health pass" given to vaccinated people and people with negative PCR tests (Polymerase Chain Reaction tests used to quickly and accurately diagnose infectious diseases, in this case Covid-19) which allowed them to travel outside of the country, to access hotels and indoor seating at bars and restaurants, as well as attend large gatherings such as concerts or festivals: "*Thus, from a health crisis, you make <u>a political crisis</u>, because with the health pass begins for France the era of parole*" (ME9).

Lastly, on November 24th 2020, president Macron mentions **climate** and **inequality crises** which pose as other battles to be won, in addition to those against terrorism and the Covid-19 epidemic: "*Our generation has together to defeat this epidemic, to confront terrorism, to confront the <u>climate crisis</u> and the <u>inequality crisis</u>*", M6).

***3.1.7 Semaine* (week):** By analyzing the concordances with the word *semaine* ("week") within our corpus, we observe that the duration of restrictions and lockdowns due to Covid-19 was

consistently broken down into weeks, rather than days or months: "*During the next 4 weeks, the rules provided by the government must continue to be respected*" (M3). The noun was found 64 times across 9 out of 15 total documents.

***3.1.8 Épidémie* (epidemic):** An **epidemic** is an occurrence of disease that is temporarily of high prevalence (Britannica, T. Editors of Encyclopedia, 2021). In our corpus, the French word *épidémie* appears 60 times across 13 documents and is used synonymously with the word **pandemic**, which is an outbreak of infectious disease that occurs over a wide geographical area and affects a significant proportion of the world's population, usually over the course of several months (Rogers, 2022). This synonymity in our corpus can be demonstrated by the following quote: "*But I will do everything I can to fight by your side, to protect you and succeed in controlling this epidemic. This ordeal, which has been going on for almost a year, is difficult for all of us.*" (M6). In this example, president Macron speaks about wanting to control the epidemic and refers to it as "*this ordeal, which has been going on for almost a year*": from this, we can conclude that he is referring to the **Covid-19 pandemic** which had, in fact, at the time of his speech on November 24th 2020, been going on for close to a year.

***3.1.9 Moment* (moment):** In our corpus, the word *moment* ("moment") appears 53 times. Sometimes, it was **moments of difficulty** which were highlighted ("*My dear compatriots, France is going through a very difficult time. No one can predict precisely how long it will last*", M2), other times, the **friendly and familiar moments** people had to give up due to Covid that were slowly returning were the focus ("*This summer of reunion and return to one's roots, this summer that will be made of friendly and family moments, will have to be nevertheless a summer of vigilance, of fight against the virus, of vaccination*", M9). When reporting on the state of France throughout 2020 and 2021, in our corpus there are several instances in which a comparison with **previous moments** was made, for example: "*But compared to the same time last year when the constraints were much stronger, we have the weapon of the vaccine, and our collective experience. And therefore, real reasons for hope*" (M10).

***3.1.10 Soir* (evening):** With the word *soir* ("evening") appearing 53 times in our corpus, we observe three main concordances: "*this* evening", "*yesterday* evening" and "*starting from …* evening". "**This evening**" is found consistently across the majority subcorpus as the speeches tended always to take place in the evening: "*My dear compatriots, this evening I want to talk to you about the coming days, about our organization in the face of the epidemic, to draw the first lessons from this crisis, and to outline in a few lines our new path*" (M4).

"**Yesterday evening**" still consistently refers to the majority leader's speeches, as proven by the next example: "*The presidential monarch, who happens to be the best virologist in France, is also the best friend of modesty, and appeared <u>last night</u> on television*" (ME8).

The last meaningful and forementioned iteration of the word "soir" ("**starting from … evening**") is used to communicate the beginning of restrictions and lockdowns: "*...the rules which are currently in force in the 19 'reinforced vigilance' departments, these rules will be extended to all the metropolitan territory starting from this Saturday evening and for 4 weeks*" (M8).