



SAPIENZA
UNIVERSITÀ DI ROMA



UNIwersYTET ŚLĄSKI
W KATOWICACH

PhD International Programme

Studies in English Literatures, Language and Translation

Double degree

Sapienza University of Rome and University of Silesia in Katowice

Taboo Language and Incest in the French and British press (2017-2022)

A CROSS-LINGUISTIC CORPUS-ASSISTED DISCOURSE ANALYSIS

Dr Sophie Eyssette

Supervisor

Prof. Margherita Dore

Co-supervisor

Prof. Andrzej Łyda

2024

Sex, for humans, is not just something we do, but also something we represent and reflect on: language plays a crucial role in shaping human sexuality and in mediating its various expressions.
(Cameron & Kulick, 2006, p. 1)

Acknowledgments

First and foremost, I would like to thank La Sapienza Università di Roma to have granted me the privilege to be their PhD Student. I am extremely grateful for the valuable freedom granted by the Department of European, American, and Intercultural Studies. I was able to carry out the topic of my choice that I investigated with great interest and pleasure over these years and to visit other universities. Furthermore, I am most obliged to La Sapienza's financial support that allowed me to study, attend conferences and stay abroad over these three years.

I am also extremely grateful to the University of Silesia in Poland for providing additional financial support. This joint PhD demonstrates how the European Union has a lot to offer for students, in terms of facilitated exchange and intercultural experiences.

Thank you to both my supervisors, Margherita Dore and Andrzej Łyda. It was a great experience to be advised by two supervisors. Margherita, thank you for your support in undertaking my research stays at Lancaster University and for letting me have the pleasure to meet Diane. Andrzej, thank you for welcoming me at the University of Silesia where I enjoyed my teaching experience.

Thank you to Donatella Montini for handling the PhD programme with so much care and efficiency.

I would like to thank Rafaella Baccolini (Bologna University) who took the time to meet me online during my Covid-impacted Erasmus Semester during my Master's final year. She was the first professor to believe that my idea could be a relevant interdisciplinary research project and pointed out the notion of taboo language that helped me shape my linguistic topic.

This PhD thesis would not be as it is without my enriching research stays at Lancaster University where I could audit stimulating classes. A special thanks to Gavin Brookes who accepted to be my additional tutor and referent at Lancaster University. It was a great pleasure to work with you.

My stays at Lancaster were full of life and support thanks to Aroa Orrequia-Barea. Our numerous conversations provided me with valuable advice and guidance. Beyond our shared linguistic interests, it was greatly enjoyable to discover the North of England with you (and share guilty pleasures for frosted cupcakes and hot chocolates topped with marshmallows!).

And of course, thank you to Diane who shared her home with me for several months over the two research stays in Lancaster. I enjoyed our endless talks about British history and political turmoil. My stays in Lancaster would not have been the same without this nice feeling of being at home, in such a charming English house.

Thank you to Rachelle Vessey for her warm welcome at Carleton University in Canada. I greatly enjoyed discussing our respective French/English research. Thank you for your time, advice, and revision.

Of course, this PhD journey would not have been manageable without the love of my family and friends. I warmly thank all my friends for providing me with emotional support. A special thanks to Sandeep for his enduring friendship and down-to-earth advice.

Thank you to my brother and my parents. A few lines are not enough to express my gratitude for listening to me, advising me and joining me on my European itinerary. Thank you for being so great!

Despite the struggles and doubts, this PhD journey was an amazing experience, enlightened by many moments of joy and intercultural discoveries.

Abstract

Taboos have been prevalent in all societies throughout history. Taboos prescribe behaviours related to death, food, and sexuality, and one of the primary sexual taboos is incest. Incest refers to the forbidden sexual relationships between family members. The incest taboo is said to be universal, i.e., proscribed in every human society. As such, incest is worth exploring from a sociolinguistic perspective as it is a social taboo, due to its prohibition in society, and a linguistic taboo, as its tabooess influences how it is discussed. This has not been previously investigated in the literature.

This study analyses taboo language related to incest, focusing on the linguistic strategies used to avoid naming the taboo. Taboo language is defined here as linguistic avoidance strategies, i.e., words that are not dared to be spoken. Therefore, this study intends to analyse the unspeakable. To do this, it addresses the methodological challenges of finding absence in a corpus. This thesis suggests a methodology for investigating discursive absence, potentially applicable to other analytical frameworks. The methodology consists in combining method and data triangulations. The method triangulation relies on corpus-assisted discourse studies (CADS), combining corpus linguistics and critical discourse analysis, and the data triangulation refers to the contrastive analysis of cross-linguistic datasets. Thus, this cross-linguistic corpus-assisted discourse analysis is applied to data extracted from the French and British press between 2017 and 2022 to investigate what is said and *not* said about the incest taboo.

The findings of this study demonstrate that: (i) even though the incest taboo is common in both cultures, the perception of the taboo radically differs between the two press coverages, (ii) the triangulations of methods and data provide more robust analyses to explore discursive absence, and (iii) the incest taboo needs to be further investigated to break the taboo of intrafamilial sexual abuse. As such, this thesis aims to offer new insights into the linguistic investigation of this silenced taboo.

Table of contents

| | |
|---|-----|
| ACKNOWLEDGMENTS | III |
| ABSTRACT | V |
| TABLE OF CONTENTS | VI |
| LIST OF TABLES | IX |
| LIST OF FIGURES | XI |
| LIST OF ABBREVIATIONS | XII |
| CHAPTER 1: INTRODUCTION | 1 |
| 1. Overview | 2 |
| 2. Objectives | 4 |
| 3. Thesis structure | 6 |
| CHAPTER 2: DEFINING INCEST | 8 |
| 1. The myth of Oedipus, a founding representation of incest | 8 |
| 1.1. The original crime of Oedipus' father | 8 |
| 1.2. Oedipus' blindness | 9 |
| 1.3. The Oedipus complex | 11 |
| 2. The definitions of incest | 12 |
| 2.1. What is a family? | 12 |
| 2.2. Incest defined in anthropology | 15 |
| 2.3. Incest defined in the Law | 17 |
| 2.4. The social implications of defining incest | 20 |
| 3. Summary | 23 |
| CHAPTER 3: TABOO LANGUAGE | 24 |
| 1. Taboo and Sex | 24 |
| 1.1. What is a taboo? | 24 |
| 1.2. Sex taboos | 27 |
| 2. Taboo and Language | 29 |
| 2.1. The critical use of euphemism | 30 |
| 2.2. Metalinguistics, vagueness and ambiguity | 32 |
| | vi |

| | |
|---|------------|
| 3. Taboo and Silence | 34 |
| 3.1. Incest, a conspiracy of silence | 35 |
| 3.2. Incest, its textual silence | 36 |
| 4. Summary | 37 |
| | |
| CHAPTER 4: RESEARCH DESIGN | 38 |
| | |
| 1. Corpus Linguistics and Critical Discourse Analysis, a method triangulation | 38 |
| 1.1. Corpus Linguistics | 38 |
| 1.2. Critical Discourse Analysis | 41 |
| 1.3. Method Triangulation | 44 |
| 2. Methodological challenges | 46 |
| 2.1. The challenges of cross-linguistic CADS | 46 |
| 2.2. Finding absence: a “dusty corner” of CADS | 50 |
| 2.3. Data triangulation, a possible methodological solution | 52 |
| 3. Corpus Building | 53 |
| 3.1. Data collection | 53 |
| 3.2. Sorting data | 56 |
| 3.3. Data distribution | 57 |
| 4. Summary | 61 |
| | |
| CHAPTER 5: THE REPRESENTATION OF THE INCEST TABOO IN THE FRENCH AND BRITISH PRESS | 62 |
| | |
| 1. A cross-linguistic analysis | 63 |
| 1.1. The topicality of each corpus | 63 |
| 1.2. The collocations of three similar lexical nouns | 68 |
| 1.3. The abortion ban in the US | 73 |
| 2. The discourse prosody of incest in the French and British press | 75 |
| 2.1. The eroticisation of incest | 75 |
| 2.2. The portrayal of incest as child sexual abuse | 84 |
| 2.3. Metaphors and paradox of the adjective <i>incestuous</i> | 93 |
| 3. The metalinguistics of the incest taboo | 99 |
| 3.1. The ambivalence of the incest taboo in the British press | 99 |
| 3.2. Breaking the taboo by saying the unspeakable | 103 |
| 3.3. Newsworthiness of the incest taboo in 2021: the diachronic analysis of the French corpus | 109 |
| 4. Summary | 113 |
| | |
| CHAPTER 6: SEARCHING FOR ABSENCE | 115 |
| | |
| 1. Developing an iterative approach | 115 |
| 1.1. The search query built on known absence | 115 |
| 1.2. The cross-linguistic challenges of semantic prosody | 118 |

| | |
|---|------------|
| 2. Optimising the search query | 121 |
| 2.1. The filtering process in the English corpus | 121 |
| 2.2. The filtering process in the French corpus | 126 |
| 3. Analysing taboo language through signalling nouns | 130 |
| 3.1. Corpus linguistics and signalling nouns | 130 |
| 3.2. A taxonomy of signalling nouns on a taboo | 135 |
| 4. Summary | 139 |
| | |
| CHAPTER 7: TABOO AND CELEBRITY, A CASE STUDY ON WOODY ALLEN AND INCEST | 140 |
| | |
| 1. Celebrity news | 140 |
| 2. The rogue celebrity | 143 |
| 3. Taboo language and celebrity | 149 |
| 4. Summary | 153 |
| | |
| CHAPTER 8: CONCLUSION | 155 |
| | |
| 1. Linguistic taboos | 155 |
| 2. Cross-linguistic considerations | 159 |
| 3. Future directions | 160 |
| 4. Concluding remarks | 161 |
| | |
| REFERENCES | 163 |

List of Tables

| | |
|---|-----|
| Table 4.1 Selected British newspapers | 54 |
| Table 4.2 Selected French newspapers | 54 |
| Table 5.1 Top twenty lexical nouns in the French corpus | 63 |
| Table 5.2 Top twenty lexical nouns in the English corpus | 64 |
| Table 5.3 Top twenty keywords of the French corpus compared to FrTenTen23 | 65 |
| Table 5.4 Top twenty keywords of the English corpus compared to EnTenTen21 | 67 |
| Table 5.5 The fifteen most frequent 5-grams in the English corpus..... | 74 |
| Table 5.6 The collocates of <i>incest</i> in the pattern <i>incest and/or...</i> in the English corpus | 76 |
| Table 5.7 Sample of concordance lines of <i>incest</i> with the collocates <i>murder, cannibalism, and suicide</i> in the pattern <i>incest and/or...</i> | 77 |
| Table 5.8 Sample of concordance lines of <i>incest</i> with the collocates <i>threat, trafficking, and danger</i> in the pattern <i>incest and/or...</i> | 78 |
| Table 5.9 The collocates of <i>inceste</i> in the pattern <i>inceste and/or...</i> in the French corpus | 79 |
| Table 5.10 Sample of concordance lines of <i>incest</i> collocating with <i>murder, cannibalism, and suicide</i> in the French corpus..... | 79 |
| Table 5.11. Sample of concordance lines with <i>incest</i> as object of the verbs <i>include, involve, and feature</i> in the English corpus | 81 |
| Table 5.12 Sample of concordance lines with <i>inceste</i> as objects of the verbs <i>dénoncer</i> (denounce), <i>évoquer</i> (mention), <i>révéler</i> (reveal), and <i>raconter</i> (tell), in the French corpus..... | 82 |
| Table 5.13 The twenty most frequent verbs with <i>incest</i> as object | 84 |
| Table 5.14 Sample of concordance lines of the verb <i>protéger</i> (protect) with <i>child</i> as object | 85 |
| Table 5.15 Sample of concordance lines of the verb <i>violer</i> (rape) with <i>child</i> as object..... | 86 |
| Table 5.16 The modifiers of <i>child</i> in the English corpus..... | 87 |
| Table 5.17 Nouns modified by <i>child</i> in the English corpus | 87 |
| Table 5.18 Sample of concordance lines of the collocation <i>child abuse</i> implying incest..... | 88 |
| Table 5.19 Sample of concordance lines of the collocation <i>child abuse</i> | 88 |
| Table 5.20 Sample of concordance lines of <i>child victim</i> with <i>incest</i> as a close collocate..... | 89 |
| Table 5.21 The concordance lines of <i>child molester, child abuser, and child rapist</i> , in the English corpus | 91 |
| Table 5.22 The ten most frequent adjectives in the English corpus..... | 91 |
| Table 5.23 The collocates of <i>sexual</i> as modifier | 91 |
| Table 5.24 Sample of concordance lines with the collocation <i>sexual relationship</i> | 92 |
| Table 5.25 The twenty most frequent nouns modified by <i>incestueux: incestuous</i> in both corpora | 93 |
| Table 5.26 Sample of concordance lines with the collocation <i>incestuous relationship</i> used metaphorically in the English corpus | 94 |
| Table 5.27 Sample of concordance lines with the collocation <i>relation incestueuse</i> used metaphorically in the French corpus..... | 94 |
| Table 5.28 Sample of concordance lines with the adjective <i>incestuous</i> used metaphorically in the English corpus..... | 95 |
| Table 5.29 Sample of concordance lines with the adjective <i>incestueux</i> used metaphorically in the French corpus..... | 96 |
| Table 5.30 Sample of concordance lines of <i>incestuous relationship</i> used literally in the British press..... | 97 |
| Table 5.31 Sample of concordance lines of <i>relation incestueuse</i> used literally in the French press..... | 98 |
| Table 5.32 Sample of concordance lines with negative semantic prosody characterising incest as a taboo.. | 100 |
| Table 5.33 Sample of concordance lines with positive semantic prosody characterising incest as a taboo... | 101 |
| Table 5.34 Sample of concordance lines showing different theories on the incest taboo..... | 102 |

| | |
|--|-----|
| Table 5.35 Sample of concordance lines where taboo refers to metalinguistic considerations | 103 |
| Table 5.36 Sample of concordance lines of the verb <i>briser</i> (break) with <i>tabou</i> as object referring to incest. | 105 |
| Table 5.37 Sample of concordance lines of the verb <i>lever</i> (lift) with <i>tabou</i> as object referring to incest | 106 |
| Table 5.38 Sample of concordance lines of <i>indicible</i> (unsayable) referring to incest | 107 |
| Table 5.39 Positive keywords in Corpus 21-22 compared to Corpus 17-20 | 110 |
| Table 5.40 The fifteen most frequent lexical nouns with their translation in each subcorpus | 111 |
| Table 5.41 Sample of the concordances <i>cette parole</i> (this speech) | 112 |
| | |
| Table 6.1 The top ten lexical nouns in the French corpus “abus AND père” | 119 |
| Table 6.2 The second search query to filter the results in the French corpus | 120 |
| Table 6.3 The final search query for automatically refining the English corpus..... | 125 |
| Table 6.4 The final search query for automatically refining the French corpus | 129 |

List of Figures

| | |
|--|-----|
| Figure 4.1 The five challenges of cross-linguistic CADS, adapted from Vessey (2013) and Taylor (2014)..... | 49 |
| Figure 4.2 Methodological framework consisting in combining two types of triangulation | 53 |
| Figure 4.3 Display of duplicate articles on Jupyter notebook | 57 |
| Figure 4.4 Number of articles in corpus EN and corpus FR | 58 |
| Figure 4.5 Number of words in corpus EN and corpus FR | 58 |
| Figure 4.6 Corpus FR breakdown per year | 59 |
| Figure 4.7 Corpus EN breakdown per year | 59 |
| Figure 4.8 The monthly breakdown of articles in the French corpus | 60 |
| Figure 4.9 The monthly breakdown of articles in the English corpus | 60 |
| Figure 4.10 Corpus FR distribution | 61 |
| Figure 4.11 Corpus EN distribution | 61 |
| | |
| Figure 5.1 <i>incest* and/or...</i> in the French and English corpora | 70 |
| Figure 5.2 Nouns modified by <i>incest/e</i> in the French and English corpora | 71 |
| Figure 5.3 Modifiers of <i>woman: femme</i> in the French and English corpora | 72 |
| Figure 5.4 Verbs with <i>woman: femme</i> as subject in the French and English corpora | 72 |
| Figure 5.5 Modifiers of <i>child: enfant</i> in the French and English corpora..... | 73 |
| Figure 5.6 Verbs with <i>incest/e</i> as object in the French and English corpora..... | 81 |
| Figure 5.7 The most frequent verbs collocating with <i>tabou</i> as object in the French corpus..... | 104 |
| | |
| Figure 6.1 Results of the close reading of the three English corpus samples | 124 |
| | |
| Figure 8.1 Conceptualisation of taboo language..... | 155 |

List of abbreviations

| | |
|-----------------------------------|------|
| Corpus-Assisted Discourse Studies | CADS |
| Critical Discourse Analysis | CDA |
| Corpus Linguistics | CL |
| Child Sexual Abuse | CSA |

| | |
|--------------------|------|
| Daily Mail | Mail |
| Daily Mirror | Mir |
| The Express | Exp |
| The Guardian | Gu |
| The Independent | Ind |
| The Telegraph | Tel |
| The Times | Tim |
| The Sun | Sun |
| La Croix | Crx |
| Le Dauphiné libéré | DL |
| La Dépêche du Midi | DM |
| Les Échos | Ech |
| L'Est républicain | Est |
| Le Figaro | Fg |
| L'Humanité | L'H |
| Libération | Lib |
| La Montagne | Mt |
| Ouest-France | O-F |
| Le Progrès | Pg |
| Sud-Ouest | S-O |
| La Voix du Nord | Vdn |

Chapter 1: Introduction

“People used to say what a wonderful father and daughter we were — they used to wipe their eyes. We were just like lovers — and then all at once we were lovers — and ten minutes after it happened I could have shot myself — except I guess I’m such a Goddamned degenerate. I didn’t have the nerve to do it.” ... “There were no consequences?” “No.” He gave one short convulsive sob and blew his nose several times. “Except now there’s plenty of consequences.”

F. Scott FITZGERALD, *Tender is the Night*, 1934

F. Scott Fitzgerald’s 1934 novel *Tender is the Night* follows Dick Diver, an American psychiatrist who falls in love with one of his patients, Nicole, at a Swiss psychiatric clinic. Nicole suffers from a nervous breakdown, and her mental neuroses are suggested to be caused by the incestuous rape committed by her father. In the lines quoted above, Nicole’s father vaguely mentions incestuous relationship with his daughter and uses the term of love to characterise it: “just like lovers — and then all at once we were lovers.” The terminology of love is characteristic of incest perpetrators who do not distinguish abuse and love in incestuous rape: they consider that their sexual assaults are signs of love and affection, or even, of sexual initiation (Dussy, 2021). Gilgun (1995) reports from her interviews with incest perpetrators that the “[m]ost striking about the perpetrators’ accounts was that almost all of them defined incest as love and care” (p. 270).

The term *incest* comes from Latin *incestus*, meaning “impure” or “unchaste.” It is derived from the prefix *-in* (“not”) and *castus* (“pure”) – that led to *caste* (“breed,” “social class”)¹. Thus, incest implies an act that goes against purity and, by extension, against the integrity of a whole. As such, incest refers to the forbidden sexual relationships between family members. The incest taboo is said to be universal, i.e., proscribed in every human society (Lévi-Strauss, 1949). In addition to its universality, the incest taboo can be perceived in two different ways. One way is to understand incest as a transgressive sexual relationship between two consenting adults from the same family. This perception of incest is widely represented in pop culture, including fiction and pornography, but is quite rare in reality (Brey, 2022; Ovidie, 2022). The second perception is to analyse incest as sexual abuse, often involving an adult abusing a child (Armstrong, 1978; Dussy, 2021; Justice & Justice, 1980). This perception of incest is often silenced, even though prevalence numbers attest to its commonness: incest occurs in one out of six families in the United States (Gilgun, 1995; Russell, 1983), 160,000 children are victim of incest every year in France (CIIVISE, 2023), and 70 % of child sexual abuse (CSA) occurs within the families in the United Kingdom (Children’s Commissioner, 2015). Thus, “although incest

¹ Oxford Dictionary of English, 2024, online version [09.09.2024], available on Apple MacIntosh

is a worldwide problem, it remains a neglected social issue. In this context the World Health Organisation classifies this problem as a *silent* health emergency” (Yildirim et al., 2014, p. 694, emphasis added). The consequences of silencing such a prevalent taboo are severe: it is estimated that the denial of sexual violence costs France approximately 9.7 billion euros for criminal justice, health expenses, lasting trauma, and an increase in borderline behaviours (e.g., alcoholism, drug addiction) (CIIVISE, 2023). To the best of my knowledge, similar estimations have not been found for other countries, but given the widespread prevalence and universality of incest, similar costs could likely be assessed. Drawing back on Nicole’s father’s reply to the doctor’s question: “there’s plenty of consequences.”

Drawing back on the extract from *Tender is the Night*, this is one rare moment when incestuous abuse is mentioned in the novel (Luong, 2010; Murphy, 1973; Stanton, 1958). Characterised by its vagueness, this extract illustrates the main point of this thesis: How incest tends to be avoided in discourse due to its tabooess. This is why this thesis aims to shed light on the linguistic taboos on the incest taboo. Therefore, silence, leading to discursive absence, is characteristic of the discourse on incest, and is the main element of inquiry in this thesis, posing methodological challenges. Silence is the act of not mentioning a topic due to its tabooess, leading to discursive absence. In other words, since a topic is silenced, it becomes absent in discourse. Furthermore, discourse can be defined as the language used in context “making sense of the world, often using language in repeated representations, narratives and arguments” (Baker, 2023, p. 237). The key term of discourse in this thesis is introduced here in a shortened definition for a general understanding, but this thesis’ key terms, along with incest, silence, and discursive absence, which are further expanded in the following chapters.

Following this definition and taking a step away from fiction that tends to represent incest in an eroticised way, as later discussed, this thesis aims to investigate the discourse on incest that may discuss or avoid the incest taboo. In other words, this thesis analyses the taboo language on incest in the sense that there are linguistic strategies to avoid naming the taboo. In doing so, this thesis aims to contribute to the literature focusing on incest (Armstrong, 1978; Brey, 2022; Durkheim, 1987; Dussy, 2022, 2021; Héritier, 2012; Justice & Justice, 1980; Lévi-Strauss, 1949; Porter & Ciba Foundation, 1984; Russell, 1983, 1999; Sacco, 2009; Wolf & Durham, 2005) but from a linguistic perspective, a focus that has been scarcely undertaken (Foubert et al., 2022; Gilgun, 1995), and, to the best of my knowledge, not cross-linguistically addressed. As such, this study investigates incest from a sociolinguistic perspective, analysing it both as a social taboo, due to its prohibition in society, and as a linguistic taboo, as its tabooess influences how it is discussed.

1. Overview

To investigate the linguistic taboos on incest, a cross-linguistic corpus-assisted discourse analysis is conducted between the French and British press from 2017 to 2022. First of all, the starting year of 2017 is selected for the extensive report of sexual assault in the press following #MeToo movement, which marks the beginning of the wider press coverage of sexual violence

on a large scale (De Benedictis et al., 2019; Tranchese, 2023). The final year of 2022 corresponds to the date when data was collected. The French press is selected as the interest in investigating the incest taboo comes from the large media discussion that took place in 2021 onwards following a high-scandal profile. In France in 2021, the publication of *La Familia grande* by Camille Kouchner revealed the incest committed by her stepfather, an eminent political scientist, Olivier Duhamel, on his stepson, Camille Kouchner's twin brother. The disclosure in the media triggered a wave of testimonies from victims of incest on Twitter under the hashtag #MeTooIncest. Inspired by the 2017 movement #MeToo, the victims of incest share their stories of sexual abuse. This wide coverage taking place in the French press is analysed and contrasted with the British press. The United Kingdom is selected due to be an English-speaking country and the place where the methodology of CADS used in this thesis has been mainly based, developed, and expanded (Baker et al., 2008; Baker, 2023; Brookes & Baker, 2021; Hunston, 2006, 2011; McEnery & Brezina, 2022; O'Keeffe & McCarthy, 2022). Finally, news discourse was deemed an interesting linguistic field of inquiry as it offers three main advantages: (i) it provides a large amount of data (ii) that can be easily collected (iii) in order to better understand how its "great potential to exert considerable influence over us" since "we all spend a lot of our time consuming it" (Bednarek & Caple, 2019, p. 7). Thus, analysing news discourse on the incest taboo may indicate how incest is socially represented.

For this purpose, an analysis in corpus-assisted discourse studies (henceforth CADS) is carried out. CADS combines corpus linguistics with an approach in (critical) discourse analysis. Firstly, corpus linguistics is a widely discussed and used linguistic methodology to investigate large amounts of data (Baker, 2023; Friginal & Hardy, 2021; O'Keeffe & McCarthy, 2022; Taylor & Marchi, 2018). The different tools, such as frequency, collocation, concordance and keywords, shed light on frequently occurring linguistic patterns. Their applications, advantages and limitations are further discussed in this thesis. These tools provide statistical founding to the analysis informed by critical discourse analysis, henceforth CDA (Ancarno, 2018; Gillings et al., 2023; Hardt-Mautner, 1995; Partington et al., 2013). CDA "seeks to highlight the structures within societies that can result in abuses of power and to consider the perspectives and voices of those who are often not regarded" (Brookes & Baker, 2021, p. 7). In this sense, this thesis examines the silence prevailing on the incest taboo. Indeed, this study is interested in investigating a widespread crime that remains silenced. However, this interest raises the methodological challenges of finding absence in discourse. The notion of absence is examined in CADS in Duguid & Partington (2018), Partington (2014), and Schröter & Taylor (2018). This thesis aims to offer methodological solutions to the challenges raised by discursive absence, as for example the risk of overinterpretation. As Sim (2007) notes, "[s]ilence is interpreted within a linguistic framework, and where there is interpretation there can always be misinterpretation, as all of us know from our experience of everyday conversation" (p. 156). For this purpose, method and data triangulations were combined to assess discursive absence. The method triangulation consists in using CADS, combining corpus linguistics and critical discourse analysis, and the data triangulation refers to the contrastive analysis of French and English datasets. Cross-linguistic CADS is an interesting field of inquiry that raises methodological challenges, outlined later in this thesis. Furthermore, this study

aims to “address the dearth of research on cross-cultural and cross-linguistic issues of sexual violence” (Vessey, 2024). Therefore, this thesis investigates how the incest taboo is discussed, or avoided, in the French and British press from 2017 to 2022. The following section introduces the research question(s) and the objectives that this thesis intends to address. The final section outlines the structure of the thesis.

2. Objectives

This thesis aims to (i) expand the notion of taboo language, (ii) develop a cross-linguistic methodology to analyse discursive absence, and (iii) contribute to breaking the incest taboo. Thus, this study addresses the main research question, along with subsidiary research questions:

- **What are the linguistic taboos on the tabooeness of incest?**
 1. How do the notions of taboo intertwine linguistically and socially?
 2. How can a silenced topic be investigated?
 3. Does a universal taboo know linguistic borders?

In addressing these research questions, this thesis strives to contribute to (i) the analytical framework of taboo language, (ii) the methodological framework of cross-linguistic corpus-assisted discourse studies, and (iii) the linguistic implications of analysing the incest taboo. Thus, this study intends to:

1. Explore the notion of taboo language

a) Expanding the definition of taboo language

The notion of taboo language can refer to swear words and insults, but this research defines taboo language as words that are not dared to be spoken, thus giving rise to linguistic avoidance strategies such as euphemisms and vague terms. This thesis aims to expand the understanding of taboo language by exploring the linguistic notions of ambiguity, vagueness, and absence. A taxonomy of taboo language is provided at the end of the analysis to serve as a conceptualising frame that other researchers can use.

b) Considering taboo language from a cross-linguistic perspective

This cross-cultural study examines the notion of taboo language from a cross-linguistic perspective, focusing on French and English. One aim of this research project is to examine whether linguistic taboos related to a universal taboo differ or converge.

c) Highlighting the interconnections of social and linguistic taboos

Linguistic taboos arise from the fear, dread, or superstitions associated with a social taboo. In this thesis, the incest taboo is examined from a social perspective: incest is prohibited and, therefore, constitutes a social taboo. Additionally, incest is a crime that is often not discussed, resulting in its silencing. Thus, the incest taboo generates linguistic taboos.

2. Examine the methodological tools for investigating a taboo topic

a) Exploring discursive absence in relation to a taboo topic

The tabooeness of the topic leads to discursive absence. As such, discursive absence should be examined from both linguistic and methodological perspectives. This cross-linguistic corpus-assisted discourse analysis addresses methodological challenges in examining discursive absence by combining method and data triangulations.

b) Considering the methodological limitations and advantages of finding absence

Various tools and analytical stages are used to investigate whether absence can be linguistically assessed. Software is applied to the data critically to address the research question regarding taboo language.

c) Contributing to CADS methodology for finding absence

This thesis builds on previous CADS research and aims to contribute to the field by integrating different approaches to studying absence and cross-linguistics. It suggests that examining absence within cross-linguistic datasets can enhance the robustness of assessments regarding what can be expected (and what cannot be found) from a discourse on taboo.

3. Bring awareness to the incest taboo

a) Contributing to the extensive literature on incest from a linguistic perspective

Extensive research on the incest taboo has been conducted in anthropology, sociology, and psychology, especially since the 1980s in the United States. However, it has been scarcely examined in linguistics. This study aims to address this gap by exploring the discursive representation of incest in the press.

b) Emphasising the definition of the incest taboo as child sexual abuse

Incest is perceived differently across cultures, but one highly taboo perception under investigation in this thesis is incest as child sexual abuse occurring within families. This understanding of incest is often silenced. This thesis investigates the linguistic consequences of silencing the widespread issue of child abuse by family members.

c) Shedding light on the cultural perceptions of the incest taboo

To the best of my knowledge, no linguistic research has examined incest from a cross-cultural perspective, calling into question the universality of the incest taboo. Thus, this study investigates cross-linguistic perceptions of the incest taboo, highlighting what is expressed and what is silenced by contrasting two representations of incest.

3. Thesis structure

This section outlines the development of the research project addressing the research questions and objectives mentioned above. It is worth mentioning that two corpora (respectively, in English and French) were built to investigate the notion of taboo language on incest. Both corpora are introduced and analysed in the following chapters, with Chapter 7 presenting a case study drawn from both datasets.

Chapter 2 – Literature Review: Defining Incest

This chapter traces the misconceptions surrounding the incest taboo back to the myth of Oedipus and its interpretation by Freud. The Oedipal complex, by emphasising the children's sexual desire for their parents, dismisses parents' responsibility and prevents perceiving incest as intrafamilial sexual abuse. Thus, this chapter emphasises the definition of incest as (child) sexual abuse within the family and introduces the other definitions of incest as universal marital prohibition and as a consensual yet transgressive sexual relationship. As such, these three definitions are expanded and examined through sociological, psychological and anthropological analyses, providing contextual insights into the understanding of incest. Furthermore, the French and British legal definitions of incest are analysed and contrasted, highlighting the cross-cultural perspectives on criminalising incest. This chapter provides the contextual background to investigate incest from a linguistic perspective in the following chapters.

Chapter 3 – Theoretical Framework: Taboo Language

This chapter outlines the theoretical framework for defining taboo language in this thesis. It explores various definitions of taboo language in linguistics before focusing on the definition of this thesis: taboo language refers to linguistic avoidance strategies. Embarrassment, politeness, and secrecy can lead to the use of taboo language that substitutes unambiguous terms with euphemisms and vague terms. Additionally, silence is examined as a crucial linguistic element in understanding the silencing of incest victims.

Chapter 4 – Methodology: Research Design

To investigate the taboo language surrounding the incest taboo, this research addresses the methodological challenges of identifying absence and conducting a cross-linguistic analysis through a research design based on method and data triangulations. Firstly, the method triangulation combines corpus linguistics with an approach informed by critical discourse analysis (CDA), known as corpus-assisted discourse studies (CADS). In addition, the challenges of conducting cross-linguistic CADS are examined. This double method and data triangulation aim to identify discursive absence. Finally, the process of building the French and English corpora is explained, and the datasets are presented.

Chapter 5 – Results Corpus 1: The representations of the incest taboo

This chapter analyses the representations of incest in the French and British press. The corpora built with the lemma *incest** in both languages are contrasted to investigate the notion of discursive absence. This chapter aims to demonstrate how the data and method triangulations can be used to assess absence using corpus linguistic tools. While the definition of incest as child sexual abuse is predominantly found in the French corpus, it appears largely absent in the English corpus. The cross-linguistic corpus-assisted discourse analysis highlights how triangulations enable identifying such absence.

Chapter 6 – Results Corpus 2: Searching for absence

This chapter offers an additional methodological perspective for examining absence. An iterative approach is developed to build corpora on a search query that deliberately avoids the term *incest* based on the hypothesis that the press may avoid using this term when reporting incestuous stories. The methodology for building the search queries in French and English is contrasted as each corpus requires a different approach due to variations in their semantic prosodies. The findings reveal that linguistic avoidance strategies contribute to ambiguity and vagueness. Furthermore, the notion of signalling noun is introduced to explore how abstract nouns contribute to the vagueness of the taboo language by euphemising taboo content.

Chapter 7 – Case study: Taboo and Celebrity

The case study provides a more fine-grained cross-linguistic analysis of taboo. It examines how the incest taboo is portrayed when a celebrity (namely director Woody Allen) is involved, exploring the characteristics of celebrity news and sensationalism. Across all the corpora, only the name of Woody Allen appears as he is accused of incestuous abuse by his adoptive daughter. This American celebrity serves as a point of comparison to investigate how the incest taboo is covered in the French and British press, particularly as references to incest become increasingly explicit. This case study shows how press coverage may shift in response to public opinion trends.

Chapter 8 – Conclusion

The conclusion addresses the research question by developing a taxonomy of the taboo language. It outlines a continuum from explicit to implicit presence, which conceptualises the findings of this thesis. Finally, the conclusion highlights the limitations of this thesis and its contributions to cross-linguistic CADS. Directions for future research are suggested.

Chapter 2: Defining incest

*O fatal wedlock, thou didst give me birth,
And, having borne me, sowed against my seed,
Mingling the blood of fathers, brothers, children,
Brides, wives and mothers, an incestuous brood.*

(Sophocles, 2020, p. 66)

This chapter seeks to shed light on the different definitions of incest, as they reflect the social perceptions of incest and have an influence on how we discuss, how we legislate or how we do research on incest. This chapter introduces anthropological, sociological, legal and literary representations of incest, but it is not exhaustive, as this study analyses linguistic taboos on incest, and not incest *per se*. It offers an introduction to incest to better comprehend the coverage of this taboo by the press in the following chapters. The first section presents a founding representation of incest through the myth of Oedipus. The second section delves into the definitions of incest in anthropology and law, followed by their implications for research and public opinion.

1. The myth of Oedipus, a founding representation of incest

The myth of Oedipus has been a source of inspiration for numerous ancient authors, including Aeschylus, Cicero, Diogenes of Sinope, Euripides, Homer, Nicomachus, Philocles, Seneca, and Sophocles (Masqueray, 1942). It is said that even Julius Caesar wrote prose on Oedipus during his moments of creative inspiration (Judet de La Combe, 2023a). Unfortunately, only fragments of these works have survived to this day, with Sophocles' tragedies being the exception. Sophocles wrote two plays on Oedipus. The first, *Oedipus Rex*, was staged in Athens around the 5th century B.C. The second, *Oedipus in Colonus*, was written twenty years later and shortly before Sophocles' death. Whereas the first play unravels the tragic fate of Oedipus, who tries to escape the god's oracle, the second play questions Oedipus' responsibility in his downfall.

1.1. The original crime of Oedipus' father

The story of Oedipus begins with his parents, King Laius and Queen Jocasta of Thebes. It is widely known that they abandon their son due to a prophecy of Delphi warning them that their child would grow up to kill his father and marry his mother. However, the reason behind the prophecy is often overlooked.

Laius is a descendant of the founders of Thebes and heir to the throne. However, his opponents constrain him to flee away from Thebes. He finds shelter by King Pelops in Argos

and is entrusted with the education of King Pelops' son, Chrysippus. Laius falls in love with the young boy. During a game, he abducts him and rapes him. It is the first case of paedophilia committed by a mortal (Judet de La Combe, 2023a).

One version of the myth says that Chrysippus decides to kill himself after the rape, prompting Pelops to curse Laius: he should either have no child, or if he has one, his child should kill him. The gods choose the second option. Another version of the myth suggests that Chrysippus is killed by his mother-in-law, Hippodamia, who seeks to secure the throne for her own children. While Chrysippus is sleeping next to Laius, she sneaks in their room, takes Laius' sword, kills Chrysippus with it, and leaves. All the elements indicate Laius as the criminal.

Nevertheless, both versions of the myth show Laius' action as reprehensible. The fact that he became involved in a forced pederastic relationship with a young boy highlights the confusing link between paedophilia and incest in the myth of Oedipus. Indeed, it is due to his father's original sin that Oedipus is doomed to commit incest. The first crime of paedophilia leads to the second crime of incest. Interestingly, this connection is still evident in modern language and media coverage, where incest, paedophilia, and pederasty are often conflated, as Allan and Burrige note: "The terms *incest*, *paedophilia* and *pederasty* are orthophemisms. The press, the public and the law generally ignore incest, except when practised as paedophilia." (2007, p. 160). Such confusion is exemplified by the Outreau case in France. In 2004, a family in the North of France was arrested for raping their children. However, the parents falsely accused their neighbours of participating in their sexual abuse, as in a prostitution network. The Outreau case received extensive media coverage and had lasting effects on the legal system. The case clearly shows the media's interest in paedophilia rather than incest, as the media was drawn to the case due to its paedophilic dimension, causing the initial case of incest to be forgotten (*L'Affaire d'Outreau*, 2023). Even today, the Outreau case is seldom remembered in France for involving incest, but instead for wrongly accusing adults of paedophilia. As such, the following chapters discuss the continuing link drawn between incest and paedophilia, especially through the vague term *child abuse* found in the data.

1.2. Oedipus' blindness

In addition, the myth of Oedipus illustrates the strength of denial and silence when humans violate a taboo. The myth begins with Laius and Jocasta having a child, despite the gods' curse. Laius, a man of impulse, forgets the prophecy during a drunken night and sleeps with Jocasta (Judet de La Combe, 2023a). They abandon the child to a shepherd whose duty is to leave the child in nature to die. The shepherd pierces the child's feet to carry him on a wooden stick; this is why the child is later named Oedipus, meaning "swollen feet." However, the shepherd takes pity on the child and decides to leave him with another shepherd he meets on the way (Sophocle, 2015). This second shepherd comes from Corinth and knows that the royal couple, Polybus and Merope, cannot have children but would love to raise a child as their own. They adopt Oedipus and raise him, providing him with a peaceful childhood and adolescence.

However, Oedipus' life takes a dramatic turn during a feast when a drunken man insults him, calling him a "bastard." This incident triggers doubts in Oedipus, and he confronts his parents,

asking them if they are truly his parents. Despite their assurance that he is their son, Oedipus remains doubtful and decides to seek answers from the oracle of Delphi, without telling them. The answer of Apollo is, as usual, cryptic, warning him he will kill his father and marry his mother. His need to understand the true meaning of the terms “mother” and “father” is left unanswered. His doubts turn into fear. Terrified by the prophecy, he decides to leave Corinth, not questioning anymore the fact that Polybus and Merope are his actual parents. He unknowingly embarks on a path that leads him to Thebes. On the way, he encounters a rude man surrounded with guards who demands that he step aside. Refusing to comply, Oedipus kills the man and his guards, unaware that the man he killed was his father, Laius. Thus, he begins to fulfil the prophecy.

Oedipus proceeds to Thebes, where he confronts and defeats the Sphinx, who has been terrorising the city. To the Sphinx’s riddle: “What walks on four feet in the morning, two in the afternoon and three at night?”, Oedipus answers correctly: “Man: as an infant, he crawls on all fours, as an adult, he walks on two legs and in old age, he uses a walking stick.” As a reward for freeing the city from the Sphinx’s tyranny, he is offered the hand of the king’s widow and the throne. Thus, he unwittingly fulfils the rest of the prophecy by marrying his mother, Jocasta, and having four children: Eteocles, Polynices, Antigone and Ismene.

Blissful years pass until the gods decide to play with Oedipus again. The city of Thebes is struck by a plague, and the people turn to their king, Oedipus, to save them once again. He consults the oracle, which, for once, is specific: the plague will end when the killer of Laius, who is living in Thebe, is found. Oedipus is puzzled as to why the murderer has not been found yet. He tries to uncover the truth and asks questions. He meets the blind prophet Tiresias, who tells him the truth, but Oedipus cannot believe him and thinks there is some political agenda in his revelation. However, Oedipus meets the shepherd who gave him to Polybus and Merope. He knows the whole truth. He tells him that Polybus and Merope are not his true parents. Jocasta, who assists to the scene, understands everything. She understands that Oedipus is the son she abandoned to a shepherd. She flees to the palace. Oedipus, who does not understand the full extent of the revelation, initially thinks that Jocasta rejects him because he is not of royal blood. The shepherd continues and tells Oedipus that he is the one who killed the former king, Laius. The shepherd was one of the guards on the day of the killing and escaped Oedipus’ assault. Oedipus understands that he fulfilled the prophecy. He goes back to the palace, where he finds Jocasta who has hanged herself. Desperate, he takes some needles from her dress and blinds himself.

The blinding act of Oedipus has been interpreted in various ways. It is seen as an act of denial, refusing to look at the infringement of a taboo (Zerubavel, 2008, p. 4). Oedipus embodies the common behaviour of looking away from a taboo: “What society expects us to ignore is often articulated in the form of strict taboos against looking, listening, and speaking” (Ibid, p. 27). Another interpretation is to see the blinding act of Oedipus as a defiant act, refusing to let the gods have the final word and wanting to speak out about the cruelty of their actions (Judet de La Combe, 2023b). Regardless of interpretation, the tragedy of Oedipus is a complex interplay of taboo, language, silence, and denial.

Furthermore, it is interesting to note that in the play, the linguistic taboo on incest already occurs. A messenger describes the scene of Oedipus blinding himself and he adds:

“He cries, ‘Unbar the doors and let all Thebes

Behold the slayer of his sire, *his mother’s*—’

That shameful word my lips may not repeat.” (Sophocles, 2020, p. 62, emphasis added)

This type of silencing is characteristic of what is still found today in the press. The social taboo generates linguistic taboos. For all these reasons, the enduring myth of Oedipus contains narrative elements that are still relevant to the discussion of incest today. The notions of taboo, denial, and silence are further investigated in the following chapters.

1.3. The Oedipus complex

The myth of Oedipus has resonated with the modern world, particularly after the Second World War with two major films, *The Infernal Machine* (Cocteau, 1963) and *Oedipus Rex* (Pasolini, 1967), as it raises questions about the free will of humankind (Sophocle, 2015). However, the myth mainly found an echo in the work of the founder of psychoanalysis, Freud. His interpretation of the Oedipus myth contributed to a misunderstanding of the incest taboo. He coined the term “Oedipus complex” to describe the repressed sexual impulses of children towards their parents (Bourdin, 2005). Yet, as noted earlier, there are multiple versions of the myth. The fact that Freud based his entire theory on the version of Sophocles raises questions about the validity of the Oedipus complex:

The hypothesis that Sophocles introduced incest into his play “Oedipus” solely for the purpose of using it as a material for his tragedy cannot be underestimated. Therefore, the Freudian interpretation, which relies solely on this text, risks masking the importance of the particular moment when the tragedy was presented to the Athenian public: in a context of fierce debates about the possibility that human beings do not depend solely on higher forces such as fate and divine forces – which can ultimately be interpreted as a debate about free will. This precision is not intended to challenge the Freudian interpretation, but rather to humbly reflect on its universality. (Cantarella & Sirignano, 2019, p. 104)

Additionally, Freud introduced the concept of the Oedipus complex as a departure from his initial theory that women’s hysteria was often caused by childhood abuse: “Sigmund Freud first heard his patients describing sexual events, he believed that the sexual abuse of children must be *common although never discussed*” (Porter & Ciba Foundation, 1984, p. xviii, emphasis added). In May 1886, at the meeting of the Society for Psychiatry and Neurology in Vienna, he defended the theory that “early abuse in childhood leads to trauma, therefore to hysteria on most women. Facing the cold reaction of his peers, he withdraws his theory to defend the Oedipal complex, more suitable for that time” (Grogan, 2016, p. 25). Therefore, it is important to note that the concept of the Oedipus complex was developed within the conservative Viennese society of the 19th century. In order to avoid controversy, Freud

abandoned his initial theory and shifted toward the Oedipus model. As Grogan notes, this allowed Freud to “eliminate the allegation that sexual abuse was so rampant in the Victorian world. It served to vindicate fathers and implicate daughters” (2016, p. 25).

Freud’s interpretation of the Oedipus myth has a significant influence. He defines the Oedipus complex as the repressed sexual desire of a child towards their parent of the opposite sex, including feelings of rivalry towards the same-sex parent. Freud believed that children’s repressed sexual feelings towards their parents were a natural part of their developmental process (Freud, 2013). “In fact, the trust of the psychoanalytic approach has been to pinpoint the child victim’s “seductive” behavior” (Brownmiller, 1993, p. 275). This “construct of the sexually fantasizing child” (Ibid.) contributes to misconceptions about incest, leading to the perception of incest as a mere fantasy today (Brey, 2022, p. 116).

Thus, the concept of the Oedipus complex raises two main issues. First, it forgets that incest is committed without the consent of the individual, as Oedipus tried to avoid the incestuous union. Second, interpreting Oedipus as a manifestation of child sexual desire perpetuates the misconception that children secretly and deeply desire their parents, thereby justifying incestuous relationships and adding to the feeling of taboo surrounding the topic. It is worth noting that Oedipus himself had no knowledge of his crime, as he searched for the killer of Laius without realising he was looking for himself (Judet de La Combe, 2023b; Sophocle, 1942). Hence, Freud’s interpretation of the Oedipus myth has left a significant mark on human consciousness.

2. The definitions of incest

Incest has traditionally been defined as a taboo within the humanities – especially in anthropology – resulting in a culture of silence on the topic. Considering incest as a taboo fuelled misconceptions that influence the legal framework, public opinion, and policies for safeguarding children. However, apprehending incest requires a clear definition of “family” in the first place.

2.1. What is a family?

The definition of family is far from reaching a consensus among ethnologists, anthropologists, or sociologists, and this thesis does not pretend to resolve this question in a few pages. However, the definition of “family” is clarified below to understand how, subsequently, incest should be understood.

In anthropology, two main types of family can be distinguished: the nuclear family, and the extended family. The nuclear family corresponds to the traditional model of wedlock, including a mother, a father, and their recognised children. In contrast, the extended family refers to polygamous families, where there are several spouses and children (Ryan & Jethá, 2023). It is worth noting that some anthropologists, and primatologists, agree to say that the model of nuclear family is culturally constructed, and that *naturally*, humans are promiscuous,

– i.e., have several sexual partners –, and are not physiologically fit for monogamous unions (Borneman, 2001; Morris, 1986; Picq & Brenot, 2012; Ryan & Jethá, 2023).

To understand how I intend to apprehend the notion of “family” in this thesis, I want to briefly go back in time to shed light on the relativity of motherhood and fatherhood. Evidence from the Pleistocene, namely when *homo sapiens* were hunter-gatherers from 2.58 million to 11.700 years ago (Johnson, 2023), shows “that the help of group members in addition to the genetic parents was absolutely essential for the survival of infants (birth to weaning) and children (weaning to nutritional independence)” (Hrdy, 2011, p. 67). Hunter-gatherers shared food, sexual favours, and the responsibility of rising children, for survival reasons (Hrdy, 2011; Ryan & Jethá, 2023). Children were not only raised by their “biological” parents, but by other members of their group, named in anthropology as “alloparents”. Alloparents contribute to a child’s survival by offering protection and food, such as “allomothers” sharing their breast milk (Hrdy, 2011, p. 180). Similarly, the notion of paternity is culturally determined. The concept of “partible paternity” refers to the beliefs in some cultures, such as in some tribes in South America, that the foetus is the product of an accumulation of sperms (Beckerman & Valentine, 2002). It implies that women have several sexual partners. Here again, by blurring the filiation, all men contribute to the protection of the child, and thus increase his/her survival chances.

By contrast, the advent of agriculture led to the model of nuclear family. With the need to labour fields, notions of property, and by extension, possession and jealousy, were introduced (Ryan & Jethá, 2023). Furthermore, the nuclear family is deeply shaped by monotheist religions, among which Catholicism, for which a heterosexual couple is *una caro*, one flesh. It implies that filiation, procreation and sexuality are bound together within a wedlock (Cadoret, 2007). The consequences of defining the family as nuclear had detrimental consequences for children. Thousands of children born outside the principle of monogamous marriage were abandoned, almost systematically resulting in their deaths due to lack of care (Ryan & Jethá, 2023).

This brief incursion in anthropology demonstrates the extreme relativity in the definition of family. Examples of single parenthood, same-sex families, or recomposed families illustrate that the nuclear family is more a traditional ideal than an actual reality – without mentioning adulteries that suggest that monogamous unions are more imposed on humans, rather than desired. In France, 66% of families can be described as “traditional”, i.e., the children live with their parents (Algava et al., 2021). By contrast, 25% are lone-parent families and 9% are recomposed families, i.e., families in which at least one child was born from a previous union. In the United Kingdom, the number of what French statistics count as “traditional” families is higher: 88% of families are married, civil-partnered, or cohabiting couples and 18% are lone-parent families (Office for National Statistics, 2024). It is worth noting that the UK statistics distinguish between married couples and those who live together without being married or in a civil union, referred to as “cohabiting” couples, whereas the statistics in France group all these statuses under the label “traditional”.

Considering this diversity of families, blood ties are not the defining factor of kinship in this thesis. Furthermore, blood filiation is challenged by new forms of kinships. Same-sex families, recomposed families, and single parenthood are mentioned above, but new ways of medically procreating children should be added to the list to understand the flexibility of defining motherhood and fatherhood. For instance, who is the mother when the child is born by a surrogate mother? The one who gave her ovule or the one who bore the child? This uncertainty is transcribed in language with euphemistic neologisms naming the situation, such as “*gestation pour autrui*” (gestational surrogacy) (Rodríguez Pedreira, 2018). Furthermore, the intricacies for same-sex couples to adopt a child show the persistent belief that nuclear family is the only “natural” way of mating, while paradoxically, medical procreation shows that a child can be conceived without a sexual relationship and without an identified father (Cadoret, 2007).

For all these reasons, I do not define a father and a mother according to blood ties, but according to the authority they have over their child. In other words, parental authority is what makes a parent. In this sense, my definition includes biological, adoptive, foster, and stepparents. In this thesis, kinship is determined by the intimacy and proximity that are shared with a child. The focus does not lie on the genitor, but on the educator. Children learn from their parents. This authority is a guidance but can also turn into abuse. If children obey their parents to lay the table, they may obey when it comes to play at a sexual “game” (Dussy, 2021). The very persons who are supposed to teach the difference between right and wrong can be the ones who purposefully blur the line, in complete (familial) confidentiality (Brey & Drouar, 2022).

This authority is the defining element of parenthood in this thesis, even though this very notion of authority is questioned, as it rises several dilemmas regarding incest. To some extent, it can be inferred that incest is possible because of the nuclear family model: “Among social mammals the male drive for sexual variety is evolution’s way of avoiding incest” (Ryan & Jethá, 2023, p. 454). Incest is possible because children’s vulnerability has been emphasised to increase their dependency on their parents and “this “naturalised” adult-child link needs to be reconsidered” (Piterbraut-Merx, 2022, p. 69). Children have been dispossessed of their autonomy within these traditional family patterns where children are completely dependent on their parents and can rarely rely on someone else’s trust. Their vulnerability is culturally constructed. This state of fact contrasts with an observation made by Malidoma Patrice Somé on the Dagaras in Burkina Faso where this author and psychologist grew up (Ryan & Jethá, 2023, p. 175). Children feel they are taken care of by the whole village, thus freely roaming around and going from one place to another. The great advantage of this is that children feel fully included, and rarely suffer from psychological problems caused by loneliness. Perhaps, beyond incest, that is the whole structure of family we should ponder on to give new autonomy to children.

2.2. Incest defined in anthropology

Anthropologists have investigated the incest taboo and given different reasons for its implementation in all societies. This section introduces seminal theories on the incest taboo, followed by a discussion on the transformation of its definition by contemporary researchers. The purpose of this chapter is not to present a comprehensive examination of every anthropological theory, but to offer a chronological overview to understand how the perception of the incest taboo has evolved over time. This overview provides contextual elements for the subsequent linguistic analysis.

Defining incest as a kinship prohibition

The first seminal theory on the incest taboo was developed by the French anthropologist Claude Lévi-Strauss in his essay *The Elementary Structures of Kinship* (Lévi-Strauss, 1949). The incest taboo, as conceptualised by Lévi-Strauss, is a universal prohibition, which means that the incest taboo is found in every society. Drawing from the linguistic rules developed by Ferdinand de Saussure (Saussure, 1916), Lévi-Strauss examines the underlying structures that shape all societies. The prohibition of incest is the founding element of his structural theory, i.e., structuralism. The incest taboo dictates that individuals must mate with members outside of their family circle. Thus, the incest taboo is defined as a kinship prohibition (Lévi-Strauss, 1949).

Furthermore, Lévi-Strauss considers incest as the passage from nature to culture. Humans leave their animalistic state to impose rules on their sexual behaviours: “It is where culture sets limits on our sexual (reproductive) activity and where sexual capability sets limits on what is culturally possible, that our humanity and our animality interact” (Leach, 1982, p. 115). The origins of the incest taboo do not stem from consanguinity reasons, but aim to reinforce the unions between tribes (Lévi-Strauss, 1949). If humans cannot mate within their own family circle, they necessarily need to create alliances outside this circle, resulting in expanding their network (Héritier, 2012). The incest taboo helps to form and maintain social order through the regulation of kinship alliances. Even though inbreeding and consanguinity are perceived as the natural reasons for the incest taboo, this conclusion is contested (Wolf & Durham, 2005). It seems that the edition of such rules —i.e., do not mate within your own family— was motivated more by cultural reasons than biological ones. The sociologist Durkheim argues that inbreeding does not lead to diseases and that the reasons for imposing the incest taboo are purely social, not biological (1987). As such, the incest taboo is perceived as a social prohibition.

Defining incest as unconsented intrafamilial sexual abuse

In contrast to Claude Lévi-Strauss’s views, the anthropologist Dorothée Dussy aims to reconsider his stance in her essay *Le Berceau des Dominations, anthropologie de l’inceste* (“The Cradle of Dominations, an anthropology of incest,” not translated), published in 2013 and reprinted in 2021 following the high-profile incest scandal in France. Dussy raises the question: If incest is prohibited, how can we explain the transgression of the prohibition? Unlike Lévi-Strauss, who views incest as a kinship prohibition, Dussy views it as a form of sexual

abuse. In her essay, Dussy describes the anthropological characteristics of incest and sheds light on “what is *said* and what is *kept silent* on incest” (2021, p. 34, emphasis in original). She does not define incest as something forbidden by social norms, but as something that occurs commonly and is kept hidden. In this sense, incest is to be understood as sexual abuse including “touching the sexual body parts of children and inducing or forcing children to touch perpetrators’ sexual body parts. These body parts include the vulva, vagina, anus, buttocks, breasts, testicles, and penis. A comprehensive definition includes non-touch incest, such as looks and talk” (Gilgun, 1995, p. 268).

Consequently, incest is an instrument of power within the family. Incestuous abuse is instigated by an older family member towards a younger one (Dussy, 2021, p. 79). The age difference leads to an asymmetrical power to which the younger family member is subjected (ibid, p. 195). The concept of “eroticised pedagogy of submission” (Dussy, 2022, 2021) implies that the child learns from early childhood what domination means, through unconsented sexuality. Children are too young to get involved in the sexual relationship with their full awareness and consent (Gilgun, 1995). By extension, the silencing process imposed on the victims by their abusers prevents children from naming what is happening to them. Silence is part of the domination mechanism since the child is constrained to remain silent on sexual abuse: “The practice of incest is protected by the absence of words to describe it” (Dussy, 2022, p. 182).

Silence is a key element in the practice of incest. Victims and incest perpetrators do not define incest with the same words when they are listened to. Interviews were conducted with convicted abusers in prison by the researchers Jane Gilgun in 1995 in the United States and by Dorothee Dussy in 2013 in France. Their results are quite similar in terms of incest perpetrators’ discourse compared to their victims. It is worth underlining that the testimonies of the convicted incest perpetrators are shaped by the legal process they went through (Dussy, 2021, p. 106), which involved reflection on their actions with lawyers, therapists, and judges (ibid, p. 68). As a result, their narratives may differ from those of the incest perpetrators who have never been arrested. Nevertheless, they avoid mentioning the incestuous dimension of their relation and describe their actions as loving and caring (Dussy, 2021; Gilgun, 1995). Gilgun even underlines her surprise to hear the incest perpetrators defining their incestuous relationships as romantic:

These were unanticipated findings. I did not hypothesize that perpetrators would view incest as caring and as romantic love. Rather, I had assumed that incest represented lack of care and, implicitly, an inability to love. It did not occur to me that perpetrators would equate incest and romance, or even incest and feelings of sexualized caring. (1995, p. 270)

Interestingly, the seriousness of the actions is downplayed by incest perpetrators both in French and English. Terms such as “make kisses,” “mistakes,” “what I have done,” “make love,” or even “sexual game,” pointed out by Dorothee Dussy, are similar to the ones listed by Jane Gilgun: “making her a better love than her mother,” “showing her I loved her,” “comforting,” “consoling,” along with the common theme of contributing to the sexual education of children with terms such as “teaching” (Gilgun, 1995, p. 272) or “sexual initiation” (Dussy, 2021, p. 80).

Both studies highlight that abusers “trivialise what is after all the narrative of a daily sexuality”(Dussy, 2021, p. 277). From the victims’ point of view, these wordings might be perceived as euphemisms to hide the truth. For instance, victims of incest never define the incestuous relationship as loving or playful, but as “abuse” (Dussy, 2021, p. 80). However, the incest perpetrators might truthfully consider their acts as caring (Gilgun, 1995, p. 172). Furthermore, abusers forget some incestuous acts, contrary to their victims, because they do not perceive the act as a traumatic incestuous behaviour, but as a trivial sexual moment (Dussy, 2021, p. 137). As such, incest is defined differently between the victims and the incest perpetrators. Euphemisms, and how terms are differently perceived as euphemistic, are discussed in Chapter 3 and analysed in the following chapters.

To conclude this section, the two main definitions of incest in anthropology mentioned above are highlighted in light of the polysemous term “taboo”, which underscores why this thesis is framed in sociolinguistics. Taboo can be synonymous with prohibition, but also with the “unspeakable.” On the one hand, there is the social dimension; on the other, the linguistic one. Lévi-Strauss defines the incest taboo as a prohibition; incest is a rule not to be crossed. Dussy defines the incest taboo as censored sexual abuse; the victims are silenced. Their perspectives differ. Dussy finds an explanation in patriarchy; the men having the dominant position do not question their role in society, and as such, Lévi-Strauss chooses to see society through the eyes of the ruler, often embodied in the man of the village (2022). Dussy admits that her stance is a feminist one. For instance, she chooses to write her essay in a more informal tone to convey the intimacy of the incestuous setting. According to her, academic language reproduces the masculinist and patriarchal model “that prohibits but admits incest, allows it to occur and favours its perpetuation, encourages silence that allows its exercise” (2021, p. 34). The challenging of patriarchy may help to contextualise other seminal writings that have exposed the prevalence of incest within societies. Notably, many female anthropologists, sociologists, and psychologists also defined incest as a widespread and hidden infringement of a well-known prohibition. The theory of the incest taboo synonymous with a kinship prohibition was dismissed to tell stories of adults who were abused in their childhood within their households, in silence (Armstrong, 1978; Justice & Justice, 1980; Gruyer et al., 1991; Gilgun, 1995; Russell, 1999).

2.3. Incest defined in the Law

The anthropological definitions of incest influence the legal framework. In France and the UK, the law has biases that differ between the two countries but tend to favour the incest perpetrator to the detriment of the victims in both cases.

In French Law

In France, the penal code defines incest as an aggravating circumstance of rape or sexual assault. This means that sexual aggression must be first established before considering its incestuous characteristics. The law, translated into English here, was revised and implemented in April 2021:

Article 222-22-3

Rape and sexual assault are considered incestuous when committed by:

1° An ascendant;

2° a brother, sister, uncle, aunt, great-uncle, great-aunt, nephew or niece;

3° The spouse, partner of one of the persons mentioned in 1° and 2° or civil union partner to one of the persons mentioned in the same 1° and 2°, if they have legal or de facto authority over the victim.

Article 222-23-2

Except in cases provided for in Article 222-23, incestuous rape is constituted by any act of sexual penetration, of whatever nature, or any oral-genital act committed by an adult on the person of a minor or committed on the perpetrator by the minor, when the adult is an ascendant or any other person mentioned in Article 222-22-3 who has legal or de facto authority over the minor.

Thus, the definition considers incest as unconsented sexual abuse between family members. However, the definition poses several issues. First, incest is indirectly targeted, hence not condemned for itself as it is an aggravated circumstance of sexual aggression (Romero, 2018, p. 1). Second, it means that sexual aggression must be first proven. However, it is difficult in the case of incest to prove the non-consent of the children (Ibid, p. 6), as the incest perpetrators are often not physically violent, but use their authority on the children to make them obey (Armstrong, 1978; Dussy, 2021; Justice & Justice, 1980; Russell, 1999). Therefore, the number of convictions is very low compared to the known number of incest victims: 800 abusers are convicted per year, while 160,000 children are known to be victims of incest every year in France (CIIVISE, 2022).

In British Law

Incest is defined in British law under the Sexual Offences Act in 1956. The definition relates to the kinship prohibition, and not to unconsented sexual abuse, and is as followed:

Incest

Incest by a man

(1) It is an offence for a man to have sexual intercourse with a woman whom he knows to be his grand-daughter, daughter, sister or mother.

(2) In the foregoing subsection "sister" includes half-sister, and for the purposes of that subsection any expression importing a relationship between two people shall be taken to apply notwithstanding that the relationship is not traced through lawful wedlock.

Incest by a woman

(1) It is an offence for a woman of the age of sixteen or over to permit a man whom she knows to be her grandfather, father, brother or son to have sexual intercourse with her by her consent.

(2) In the foregoing subsection "brother" includes half-brother, and for the purposes of that subsection any expression importing a relationship between two people shall be taken to apply notwithstanding that the relationship is not traced through lawful wedlock.

This legal definition of incest is marked by heteronormativity and patriarchy. Heteronormativity because each gender is associated with the opposite gender. Such definition is problematic as incest can involve two persons of same sex (Dussy, 2021; Elliott, 1993; MacMullen, 1990). Hence, it questions what penalty can apply if this situation is not acknowledged in the law. Patriarchy because the definition of "incest by a woman" depicts her in a passive role where "she *permits* a man ... to have sexual intercourse with her *by her consent*," whereas the definition of "incest by a man" depicts him in the active role of the investigator: "to *have* sexual intercourse." This definition is also problematic because it does not acknowledge that women can rape (Elliott, 1993) and men can be raped, even though "in male-dominated cultures, males do not want to accept that they can be and are victims of rape" (MacMullen, 1990, p. 9). This acknowledgment is even more taboo, which leads men who have been sexually assaulted by a woman to lie about the gender of their aggressor, which tends to underestimate the number of female abusers (Elliott, 1993; Hetherington, 1999). Furthermore, "father-son incest is thought to be even more underdiagnosed than the other types of sexual abuse" (Yildirim et al., 2014, p. 696), where the incest taboo is paired with a homosexual denial. The limited scenarios considered by the law prevent children from being protected: "children will continue to be abused because we have decided to listen to them only in an established framework" (Elliott, 1993, p. 14).

Thus, the definitions of incest in the French and British laws differ. The French law criminalises incest as sexual abuse, while the British law forbids any willing incestuous union. As such, the British law specifies an age of consent for women: "for a woman of the age of *sixteen or over*" – but not for men. Therefore, it suggests that (i) incest is not considered child abuse and, (ii) the law implies that women are necessarily younger than the male offenders. Indeed, it is an offence for the man to be with his "granddaughter," but the equivalent for the woman is not "grandson" but "grandfather," implying that she is necessarily younger than the man (or that the woman could *not* be sexually involved with a younger child). Thus, incest is forbidden in the UK when two *consenting* adults are in an incestuous relationship. The notion of sexual abuse is not retained. However, the Sexual Offences Act 2003 considers two categories involving sexual abuse within the family: the familial child sex offences, including "sexual activity with a child family member" and "inciting a child family member to engage in sexual activity," and sex with an adult relative, including "sex with an adult relative: penetration" and "consenting to penetration." In addition, the family relationships specified include biological parents, foster parents, stepparents, and adoptive parents. Referring back to the definition of family provided above, the relationship is not defined by blood but by a shared household where the parent "is or has been regularly involved in caring for, training, supervising or being in sole charge of" the child. Interestingly, British law does not define this intrafamilial sexual abuse as incest, but as a type of sexual offences. This difference in definitions influences how the prevalence of incest is measured in each country, as shown in the following section.

2.4. The social implications of defining incest

The prevalence of incest in France and the United Kingdom

In France, the Commission of Inquiry CIIVISE, that stands for Commission Indépendante sur l’Inceste et les Violences Sexuelles faites aux Enfants (Independent Commission on Incest and Sexual Violence against Children), estimates that 160,000 children are victims of incest every year. In contrast, 800 abusers are convicted per year (CIIVISE, 2022). This Commission, implemented by the French government for a two-year mandate, addresses this discrepancy between the number of victims and the number of incest perpetrators by tackling the legal failures. Furthermore, not believing children has consequences that can be calculated. It is estimated that child abusers cost 9.7 billion euros every year in France, in terms of long-life impacts on victims’ physical and mental health (CIIVISE, 2023). Other interesting numbers given by CIIVISE refer to the overarching secrecy around incest. It is estimated that one out of two victims speaks out after ten years, one out of two victims does not dare to speak about it because of shame, guilt, and fear of family’s reaction, and four out of five families do not protect the child who speaks out. The incest taboo is defined here as unconsented sexual abuse that is unspeakable.

In the UK, there is an absence of official data on incest. For instance, incest is not classified as one category but is subsumed into the “sexual assault” category in the latest report on sexual offences in England and Wales by the Office for National Statistics (2021). Incest is classified in the “other sexual offences” section, along with “sexual exploitation of children” and “sexual grooming.” Similarly, a report on crime statistics of 2011 mentions “incest” only once and it is in an annex about crime classifications (Government Statistical Service, 2011). During this research, only one report by the Children’s Commissioner for England organisation was found, providing some estimations. According to this 2015 report, 70% of child sexual abuse occur within the family environment, and 34,000 children were known to be victims of incest from 2012 to 2014. Nevertheless, the scarcity of national data reveals the uneasiness regarding this topic. Brownmiller made a similar observation in the 1990s about “the national incidence of child molestation” in the United States:

Is the sexual abuse of minors extreme and aberrant behavior committed by a small number of unfortunates who cannot help themselves – and whose crimes have been blown out of proportion by the tabloids to such an extent that they have ticked off *hallucinating fantasies in “hysterical” mothers and their “lying” children?* Or is it an all-too-real and *rather common* experience? The FBI’s Uniform Crime Reports are no help to us here, for incredible as it sounds, although they can tell us all about the theft of automobiles, the government’s crime fighters have *never produced* a national analysis of sex crimes committed against the defenseless young. (1993, p. 272, emphases added)

More than 30 years later, a similar conclusion can be drawn in both countries. Many mothers and children are still considered to be lying when they report child molestation, and even today, very few official numbers are provided to estimate the prevalence of incest. This further illustrates how taboo leads to silence, including in research.

The incest taboo in research

As mentioned above, incest is an avoided topic, surrounded by silence. People prefer to look away when they are confronted with a taboo (Zerubavel, 2008). As such, the academic field is also impacted by secrecy, because “what we ignore or avoid socially is often also ignored or avoided academically” (Ibid., p. 13). Incest, and by extension, child sexual abuse, is challenging to investigate because secrecy still prevails (Sariola & Uutela, 1996; Pereda et al., 2009; Yildirim et al., 2014; Montgomery et al., 2015). Therefore, the incest prevalence is higher than what is officially reported. There is a need to shed light on incest as a sexual crime as Yildirim et al. point out in their study: “Incest is a taboo and a neglected social problem in all cultures. We aimed to focus attention on this important form of sexual abuse” (2014, p. 697). Similarly, Pereda et al. estimate that the “experience of child sexual abuse is a problem of considerable magnitude in all the societies analysed” (2009, p. 697). Indeed, incest as sexual abuse is underreported, and therefore underdiagnosed. A report from the World Health Organisation (WHO) published in 2004 has a telling title: “Child Sexual Abuse: A Silent Health Emergency”, where “silent” is the key word:

CSA is a *silent* health emergency. It goes *unnoticed*, is grossly *under-reported* and poorly managed. It is *surrounded by a culture of silence and stigma*, especially when it occurs *within the sanctuary of the home by someone the child knows and trusts*. The main responsibility for protecting the child lies with the family. However, the perpetrator may be a family member, close family friend or an adult in a position of authority over the child. (Regional Office For Africa, 2004, p. 1, emphases added)

The focus of the report is on the African Region, even though WHO estimate that globally 40 million children aged 0-14 years endure various types of abuse and neglect, with the most severe being child sexual abuse. Explanations for these appalling numbers may be multifactorial:

Factors such as the *secrecy* which surrounds the abusive situation, the *shame* felt by the victim when speaking about what has happened, the *criminal penalties* to which the abuser may be subject, and the *young age of victims* combined with their *dependence upon adults* mean that very few victims come forward at the time of abuse, it therefore being highly probable that official statistics underestimate the true extent of the problem. (Pereda et al., 2009, p. 334, emphases added)

As a result, the main consequence of under-diagnosed incest is not to believe children when they testify to be victims of incest. Since incest is supposed to be prohibited, in the sense that the incest taboo is a marital prohibition, incest is not believed when it occurs as intrafamilial sexual abuse (Armstrong, 1978; Dussy, 2021; Justice & Justice, 1980). Children are considered untrustworthy and the theory of the syndrome of parental alienation, developed by the American doctor Richard Gardner in 1987, largely contributed to it. This theory argues that the child is manipulated by one parent in retaliation against the other parent (Gardner, 1987). Almost systematically, the manipulative parent is said to be the mother, while the falsely accused parent is the father. Children are not believed to be abused, but to be alienated (Barnett, 2020). This way, their narratives are dismissed.

The theory of parental alienation has a lasting impact in the legal system, as a report from the British Ministry of Justice states: “Others noted that professionals were too ready to see signs of alienation, and so silencing children, rather than assessing further what the child may have witnessed or experienced” (Hunter et al., 2020, p. 79). The tendency to doubt children’s testimonies still prevails in research. For instance, a research article mentions that the “risk assessment” is to consider that some children who testify to be victims of sexual abuse may lie (Runyan, 1998, p. 494). On the contrary, some researchers point out that the risk of having lying children is minimal, and that there are more children who are not believed, despite the truthfulness of their abuse (Armstrong, 1978; Brey & Drouar, 2022; Dussy, 2021; Romero, 2020; Rougier, 2020; Russell, 1999). The syndrome of parental alienation was later criticised for promoting a patriarchal view of society where men’s actions are not questioned. It was later said that Gardner was a masculinist and had no proof for his theory, which is recognised neither by the American Psychiatric Association nor the American Medical Association (Gumbel, 2003; Oliver, 2003).

Furthermore, the common belief that children are “passive and vulnerable” leads to question “their testimony, and the low value of truth given to it” (Piterbraut-Merx, 2022, p. 80). The dependency of children on adults is construed through the idea that children are vulnerable and gives reasons for adults to have rights over them. However, a child’s vulnerability is not naturally given, but socially constructed (Piterbraut-Merx, 2022, p. 75). Children are deprived of any autonomy or legal rights and are entirely dependent on their parents (Gilgun, 1995, p. 267). It partly explains why children’s testimonies are often disbelieved and why there are scarce legal actions to find protection away from their parents (Armstrong, 1978; Dussy, 2021; Brey & Drouar, 2022). When incest is discovered or denounced, it often occurs when incest has stopped and the child is a teenager or young adult (Dussy, 2021). As such, this intrafamilial sexual abuse remains unspoken, even in fiction, where the depiction of incest highlights the transgression between two consenting adults.

The eroticisation of the incest taboo in fiction

Novels, fairy tales, movies, and series, often represent incest as a consensual sexual relationship, despite being depicted as transgressive. For example, the popular series *Game of Thrones* (2011-2019), created by David Benioff and D. B. Weiss, portrays a sibling incest between two twins, even though this type of relationship has never been scientifically found, as sibling incest is initiated by the older sibling towards the younger one (Dussy, 2021; Justice & Justice, 1980). Such depictions are unreliable and can spread misconceptions about incest worldwide by portraying it as a relationship between two consenting adults. Films have portrayed incest in a consensual and even desirable manner in movies such as *The Game is Over* (1996), *Stepfather* (1980), *The Prodigal Daughter* (1981), and *My Mother* (2004), to name just a few. However, these representations omit the violence inherent in incest (Brey, 2022). Even though it is not the aim of fiction to mirror reality, these films tend to create an eroticised perception of incest that can mislead public opinion. Nevertheless, some movies represent incest differently, such as *The Celebration* (Vinterberg, 1998), a Danish movie that won the Jury Prize at the Cannes Film Festival. The film depicts a man who confronts his family at a

dinner party and exposes the incestuous abuse he suffered at the hands of his father. Despite the revelation, the guests remain in denial and stay silent. Such films highlight the harmful effects of incest and may challenge public opinion about the taboo subject, but they are scarce.

Furthermore, incest is eroticised in pornography where figures of the MILF – Mother I'd Like to Fuck, and by extension, of the stepmum, participate in picturing consensual incest. Whereas the MILF is not the mother, but has all her characteristics, and is often the mother of one's best friend, the step-mom is a sexy mother with whom the step-son sleeps (Ovidie, 2022). Then, from the terms *stepmum/stepdad*, pornography offered videos with "just *dad*. From the stay-at-home mum in high heels and stockings who arouses friends' fantasies, we slid towards the (step-)father who rapes his (step-)daughter" (Ibid, p. 130). The representations of incestuous rapes contribute to child pornography: "Whereas in the stepmum genre, the actress is usually only about ten years older than the stepson actor, in father and daughter videos, the age gap is fully embraced, with men being two to three times the age of young teen women" (Ibid, p. 131). Therefore, even though there are no regulations prohibiting such pornographic depictions, it is worth considering the potential consequences of these videos portraying incest both in consensual and abusive ways. The latest report published by the French Senate on pornography describes that just on Pornhub —one of the most popular pornography sites— 71,608 videos promote incest and paedophilia, with titles such as "daddy fucks his teen daughter" (Billon et al., 2022, p. 88). On the six most mainstream porn platforms, categories referring to incest are entitled: "family," "family fantasy," "family secret," "family therapy," "disturbed family," "fuck mummy," "taboo mother," "mom and son," "stepdaughter," "brother and sister," "step fantasy" (Ibid., p. 89). As can be noticed, incest and all different types of kinships are eroticized, even though the terms "fantasy," "secret," or "taboo" denote the awareness that such likings are transgressive.

3. Summary

This chapter provides a contextual framework to comprehend incest in the following analysis. As observed, incest is either defined as marital prohibition, intrafamilial child sexual abuse, or consented sexual relationship. These different definitions emphasise either the prohibition or the secrecy or the transgression of the taboo. These perceptions of incest are later found in the French and British press and are discussed in Chapters 5, 6, and 7. The following chapter, Chapter 3, emphasises the notion of taboo language, thus introducing the theoretical framework of this thesis.

Chapter 3: Taboo Language

When speakers refrain from talking about the taboo, they empower the taboo.
(Jay, 2000, p. 128)

In the previous chapter, different definitions of incest were examined as well as their consequences on public opinion, in legal framework and research. In the present chapter, the aim is to demonstrate the connections between a social taboo and its linguistic taboos. The link between a social taboo and its linguistic taboos is relatively underexplored in linguistics (Allan, 2019; Allan & Burrige, 2006; Crespo-Fernández, 2015, 2018; López Díaz, 2018a; Pedraza, 2018; Popescu, 2018a; Sandoz, 2022), which could explain the various definitions of taboo language in linguistics (O'Driscoll, 2020). Furthermore, incest is an understudied research topic due to its tabooess (Sariola & Uutela, 1996; Pereda et al., 2009; Yildirim et al., 2014; Montgomery et al., 2015). In addition, linguistic research on incest is very scarce (Foubert et al., 2022) and no studies have attempted to investigate comprehensively the link between social and linguistic taboos on incest. Taboos are described here through their social causes and linguistic consequences. Thus, this chapter examines the social and linguistic manifestations of the incest taboo. The first section introduces a comprehensive overview of taboo and focuses on the taboos related to sexuality. Then, the notion of taboo language is expanded to include euphemism and vagueness leading to ambiguity. Finally, the silence surrounding incest is examined.

1. Taboo and Sex

1.1. What is a taboo?

Fear and coercive control

“Taboo” is a Polynesian term and was introduced to Indo-European languages and cultures in 1777, when Captain James Cook noted in his logbook that some rituals or forbidden food were called taboo (Allan & Burrige, 2006; López Díaz, 2018a; Sandoz, 2022). In Tongan, “tabu” means “sacred” or “holy,” thus referring to a forbidden act (Brandes, 2018). Anthropologists took a sharp interest in the concept of taboo across cultures, and observed that taboos serve as coercive tools within populations (Webster, 1973):

[I]n order to be called a taboo, violations of the rules must be met by sanctions. Sanctions might simply take the form of a verbal rebuke, or might result in more serious consequences such as ostracism or -in extreme cases- incarceration or death. Taboos, then, constitute an essential element within any system of social control. (Brandes, 2018, p. 374)

The fear of defying the forbidding rules leads to social coercion and thus, social cohesion. Taboos proscribe behaviours considered detrimental to a social group (Kacprzak, 2018) to preserve its well-being and social order (Freud, 1913). Therefore, taboos are a fundamental

and structural element in every culture (Lévi-Strauss, 1949). They encompass restrictions not only on behaviours, gestures, and rituals, but also on language. Social taboos give rise to linguistic taboos (Diffloth, 2014; Freud, 1913). Linguistic taboos emerge because social taboos impose “restrictions” on language (Widłak, 1965, p. 945). This is why taboos were initially investigated in anthropology and ethnology before becoming a subject of interest in linguistics (Sandoz, 2022).

Furthermore, the fear of danger gives rise to taboos, uniting a population in a common dread (Freud, 1913). This fear is transcribed into language where terms are purposefully banned and avoided. As Jay (2000) explains, “taboos are sanctions on thoughts and behaviours that a society finds too powerful, dangerous, or mysterious to consider openly. Threatening thoughts and words are forbidden and avoided” (p. 193). This is the word magic, namely the belief that words have a real influence on the world, or what Austin (1991) terms “the performativity of language,” by showing that words have the power to bring about real-world consequences. This avoidance of naming the feared object can be illustrated by a linguistic study in the early 20th century pointing out that the word *bear* was avoided in Slavic, Baltic and Germanic languages (Meillet, 1906). It was replaced by periphrases, leading Meillet to suppose that a taboo may have contributed to the disappearance of the word *bear* in those languages. There was neither a linguistic cause, as the word was not complex, nor an external cause, as bears were present in Slavic, Baltic and Germanic territories. However, the fear associated with dangerous animals led people to avoid naming them directly, choosing derivatives instead. Interestingly, similar examples of linguistic taboos related to the fear of animals can be found in other geographic areas, like in south-eastern Papua New Guinea where fishermen avoid naming dangerous creatures like sharks or crocodiles “for fear of inviting attack” (Allan & Burridge, 2006, p. 39). Fear of sanctions, or fear of danger, creates linguistic taboos in both cases. They can be found universally, even though they are culture-specific as it is “a common communicative code within the speech community” (Brandes, 2018, p. 389). Taboos also change over time (Khairullina et al., 2020; Popescu, 2018a). Social and linguistic taboos are shaped by diachronic and synchronic variations, as explained in the following section.

Diachronic and synchronic variations

Diachronic variation of taboo refers to the evolving status of taboos within the same society. Widłak considers that “primitive societies” do not have the same taboos as “modern societies” (1965, p. 935). He suggests that taboos in primitive societies are determined by religion and superstitions, while those in modern societies are based on social norms and ethics. Nevertheless, taboos continue to exist, albeit motivated by other reasons. Fear has been replaced in modern societies by “norms of social ethics” that continue to forbid some practices and words without further explanation (Widłak, 1965, p. 934). Shame and embarrassment are the modern fears (Sabri et al. 2010). Thus, societies abandon some taboos in favour of new ones (Sablayrolles, 2017) as “many categorical imperatives are de-tabooed, while others are tabooed” (Dyoniziak, 2018, p. 109). For instance, the fear of dangerous animals has evolved over time. Conversely, old age has become a new taboo in modern cultures. While in the past, old people were considered for their wisdom, today they are often

marginalised in facilities with euphemistic names that emphasise rest and peace, such as “retirement home” or “nursing home” (Kacprzak, 2018). Similarly, old people are now referred to as “60+” and “80+,” or as “elderly” rather than “old.”

Likewise, the perception of incest has undergone diachronic variation. The term *incest* was tabooed and avoided in court when judging cases of incest as sexual abuse in the 19th century. In a trial in 1845, incest was referred to as “a crime ‘that one blushes to name’” (Ambroise-Rendu, 2016, p. 61). Such a misname is perceived as censorship that hinders the acknowledgment of incest and the legal response to it. On a larger scale, in the 1970s and 1980s, both in France and the UK, having sex with children was not condemned but rather encouraged. In Parisian intellectual circles, paedophilia was accepted in the name of child sexual freedom (Verdrager, 2021). In 1977, prominent intellectuals like Foucault, Sartre, Barthes, de Beauvoir and 66 others signed a petition to defend three men arrested for sexually abusing teenagers under the age of 15. The Parisian intelligentsia found it disproportionate to classify as a “crime” the act of having sexual intercourse with underage boys and girls and photographing their “sexual games” (*À Propos d’un Procès*, 1977). Similarly, the lobby of Paedophile-PIE-group in the UK was still active in the 1980s, as well as the pro-incest lobby in the US (Porter & Ciba Foundation, 1984). Those groups promoted “the ‘sexual liberation’ of children whom they see as being denied their right to full sexual satisfaction” (Porter & Ciba Foundation, 1984, p. xviii). However, perceptions of incest and paedophilia have gradually changed. In France, in 1986, the first victim of incest spoke out on television. Eva Thomas wrote a book recounting the rape committed by her father (Ambroise-Rendu, 2016). Likewise, paedophilia is also no longer accepted as it used to be and is harshly criminalised (Verdrager, 2021), and it is visible in language. In French, the neologism *pédocriminalité* (paedocriminality) is gradually preferred and used over the term *pédophilie* (paedophilia), because the etymology of *paedophilia* literally means “loving children” and could be positively used in some circles. However, the suffix “criminality” emphasises the transgression (Verdrager, 2021). This linguistic trend is supported by the data from the French press, where *pédocriminalité* is a slightly more frequent term than *pédophilie* (see Chapter 5).

Synchronic variation of a taboo refers to geographic or cultural differences in perceiving a taboo within the same time period. Social taboos can vary from one country to another, including their linguistic manifestations, because societies do not share the same prohibitions (Kacprzak, 2018). For instance, a teacher in the United States was accused of promoting pornography in her class by showing Michelangelo’s statue of David. The same teacher was later invited by the Mayor of Florence in Italy when he learned that she had been dismissed for showing a masterpiece of his city (Garau, 2023; Perilli, 2023). The perception of nudity, whether it is seen as heroic, erotic, or pornographic, differs from one society to another. As Leach (1982) notes, “there will then be taboos against appearing ‘naked’ in inappropriate situations, though just what counts as either naked or inappropriate varies greatly from place to place for no very obvious or logical reasons” (p. 117).

Therefore, Allan and Burridge argue that “there is no such thing as an absolute taboo (one that holds for all worlds, times and contexts)” (2006, p. 11). Regarding incest, they contest its universality, stating, “[y]ou may believe it taboo for an adult to have sex with a minor, but hundreds of thousands of people have not shared that taboo, or else they have put it aside. Incest is tabooed in most communities, but Pharaoh Ramses II (fl. 1279-1213 BCE) married several of his daughters” (Allan & Burridge, 2006, p. 10). As a matter of fact, the practice of royal families to mate within their own family was a sign of distinction, differentiating them from common mortals. In the case of Pharaoh Ramses II, he represented God and the union of Isis and Osiris, who were brother and sister, and lovers. Thus, in the name of divine right, the Pharaoh was allowed to marry his sister (Guibal, 2010).

However, the universality of certain taboos can be argued. Freud considers that there are three universal taboos: cannibalism, murder, and incest (Hirsch, 2020). These fundamental prohibitions structure every culture as they ensure their survival. Thus, even though Allen and Burridge (2006) contest the universality of some taboos, they acknowledge that every society shares taboos related to bodies, bodily fluids, organs, sexual acts, diseases, death, killing, and food. The specific forms of those taboos vary from one culture to another, but examples can be found in every culture, influenced by traditions and religions. Likewise, sex has been universally tabooed, but the “ways of understanding and organizing it are not universal, they are variable across cultures and through history” (Cameron & Kulick, 2006, p. 3). Thus, the next section explores taboos on sex in particular, to later understand taboos on incest.

1.2. Sex taboos

Tabooed body

In the preface of *The Vagina Monologues*, Eve Ensler (2004) explains that the word *vagina* was censored in some countries where her book was published with the simplified title *The Monologues*. The author is mesmerised by the taboo surrounding the female sexual organ which even leads to reject the scientific term *vagina*. Over time and across cultures, the human body and its sexuality have been tabooed and constrained (Ryan & Jethá, 2021). For instance, the natural process of women’s menstruation, a sign of life, has been tabooed in various religions and associated with impurity and dirtiness, condemning women to be set aside during their periods (Allan & Burridge, 2006). It can be argued that the sexual body only becomes dirty when consider as such (Douglas, 2005; Leach, 1982).

Therefore, sexuality is “surrounded by a mystery almost sacred” (Widłak, 1965, p. 935), and language reflects the resulting discomfort. From a young age, parents teach their children to use “child language” to name their body, like *tushy* or *poo-poo* (Jay, 2000, p. 127), in other words, “[p]arents tell their children not to use dirty words” (Jay, 1992, p. 1). The term “dirty word” itself shows the cognitive connection established between the body and dirtiness. Children are taught to view their body as dirty and are sanctioned whenever they use the wrong terms. Consequently, humans learn from their early childhood how to speak and *not* to speak about sex (Jay, 2000). Even families who define themselves as “open-minded” and

claim to be able to “discuss any topics” find it challenging to talk about sexuality and the changing bodies of preteens (Vinel, 2017, p. 3). Speaking about sexuality often leads to “avoidance behaviours” from parents who tend to use euphemisms to avoid direct conversations (Bennett & Harden, 2019, p. 7). Language plays a significant role in shaping our understanding of sexuality. Thus, “the child who is told that “sex is dirty and sinful” will thereby learn a negative emotional response to sexual stimuli” (Staats, 1996, p. 83). In the same way, the term *vagina* is considered taboo despite being a medical term, as the scientific terms are rarely used in everyday conversations. Instead, substitute expressions are common, “because sex has been a taboo topic for so long, many euphemisms have developed to describe genitals, sexual acts, sexual body parts, and body products” (Jay, 2000, p. 128).

As a result, the taboos on sex lead to “a great deal of verbal play and figurative language” (Allan & Burrige, 2006, p. 144). Social embarrassment regarding sexuality, including sexual organs and acts, is expressed through various euphemisms or dysphemisms (Allan & Burrige, 2006; Crespo-Fernández, 2015, 2018). For instance, in the 18th century, masturbation was believed to cause madness, weakness, deformities and diseases. The church named this “abomination” as “the horrible act,” “the solitary/destructive vice” or “self-abuse” (Allan & Burrige, 2006, p. 146). Today, the taboo around “masturbation” persists through slang expressions like “beat the bishop”, “finger” or “jerk off” (Ibid., p. 147).

Taboo words and taboo language

Taboo language may include a long list of taboo words referring to animals, women, stupidity, racism, or homophobia (McEnery, 2006), but according to offensiveness ratings, the most offensive words relate to sex. For instance, Jay and Janschewitz (2008) note that the top-rated offensive words are “*cocksucker, cunt and fuck*” (p. 277). Similarly, Bednarek observes that the words “*fuck, shit, cunt, motherfucker* are consistently rated as the most offensive, while words like *hell* or *damn* are considered mild” (2019, p. 2). Furthermore, Allan and Burrige (1991) note that bodily effluvia like shit, vomit, sperm, urine and menstrual blood receive the highest revoltingness ratings in their survey, while nail parings, breath, blood from a wound, hair clippings and breast milk were considered the less revolting. They suggest that since the organs “used in sexual intercourse are also used in urination,” it conceptually links sex and dirtiness (Allan & Burrige, 1991, p. 54). For instance, the word *fuck* is perceived as one of the most offensive taboo words in English (McEnery & Xiao, 2004; McEnery, 2006; O’Driscoll, 2020). Allan and Burrige expected “that the ratings will vary from society to society, but the taboo topics are almost universal” (1991, p. 52). Various aspects of the body including its effluvia and sexuality have been tabooed and “it is this power of dirty words that leads them being condemned and their users censured” (Allan & Burrige, 2006, p. 42). It explains why such taboo words are still censured today on some national televisions, like in the United States (Bednarek, 2019).

Nevertheless, a distinction is drawn between taboo language and taboo words in this thesis. It has been noticed that taboo language is repeatedly defined as forbidden words (Allan & Burrige, 2006; Rosewarne, 2013), bad language, or swear words (Bednarek, 2019; Vallery &

Lemmens, 2021). O'Driscoll (2020) found various terms in the titles of linguistic journals that show the diversity and flexibility of the term “taboo language”: “swearing”, “taboo [word/terms/lexis]”, “[verbal/linguistic/word] taboo”, “obscenity/obscene language/obscene words”, “curse/cursing/curses/curse words”, “offensive language/offensiveness”, “profane language/profanity”, “dirty word(s)”, “vulgar language”. However, this thesis defines “taboo language” as the words that are purposefully avoided (Ambroise-Rendu, 2016; Goudaillier, 2018; Kacprzak, 2018; López Díaz, 2014; Popescu, 2018b; Sandoz, 2022). In this sense, taboo language refers to linguistic avoidance strategies. Taboo language may be used to avoid naming a taboo because of embarrassment, fear or censorship. For instance, this thesis is interested in exploring whether the term *incest* is sometimes avoided, because tabooed, and is substituted with other terms that imply incest indirectly (see Chapter 6). It has already been noted that victims of incest testifying in French court use linguistic avoidance strategies (Romero, 2020). Instead of clearly naming the type of abuse, the victims use expressions like “he did dirty things to me” or “filthy stuff” (Ibid., p. 4). This finding shows two things: (i) It confirms that sexual taboos are assimilated to dirtiness, and (ii) the victims themselves avoid naming incestuous abuse. In the data analysed in Chapter 5, some testimonies of victims in court are reported by journalists who also avoid describing incestuous abuse in clear terms. Instead, the notion of “unspeakable” is emphasised to name incest. Thus, taboo language is to be differentiated from swearing words in this thesis. It is worth noting that there are no swear words referring to incest in particular, except perhaps insults against one’s own parents. Furthermore, this study is not looking into dialogues or oral speech, but written ones, where swear words are rarer (Bednarek, 2019). The following section explores the linguistic avoidance strategies of taboo language that include euphemisms and vagueness leading to ambiguity.

2. Taboo and Language

The paradox of taboos is their “noticeable presence” (Rosewarne, 2013, p. viii), because even though they are not supposed to be discussed, “we keep talking about [them], pretending not to do so” (Popescu, 2018a, p. 144). This paradoxical nature of taboo draws from the Freudian theory of emotional ambivalence (Crespo-Fernández, 2015). Freud underlines the ambivalence that humans have with a taboo. Even though they are fully aware of the prohibition, they are drawn to defy it. This temptation to violate a prohibition is repressed in the subconscious (Freud, 1913). Similarly, Freud considers that the incest taboo is repressed by the child and then expressed through sexual desire for their parents, which he names the Oedipal complex (see Chapter 2). Therefore, individuals are both afraid of violating the taboo and attracted to violating it. Likewise, even though individuals are taught not to use taboo words or discuss taboo topics, they are drawn to them. Furthermore, taboo topics cannot be avoided, which results in linguistic avoidance strategies. As Jay (2000) notes, “[m]odern taboos on references to death, sex, excretion, religion, and disease require speakers to avoid these topics, and euphemisms and circumlocutions are constructed to talk about these facets of life” (p. 197). Avoidance linguistic strategies, such as euphemisms, are used to discuss taboos

indirectly. It is therefore important to consider their possible categorisation as expounded in the literature.

2.1. The critical use of euphemism

Typology and limits

As discussed above, in this thesis, taboo language is not defined as swear words, but as linguistic avoidance strategies. In this sense, taboo language includes euphemisms – words or expressions used to soften the emotional impact of a dispreferred term due to its sensitive or taboo nature. The term *euphemism* originates from the Greek word *euphēmismos*, “use auspicious words,” with the prefix *eu-*, “good,” and *phēmē*, “speech” (Payne, 2023). The modern meaning of *euphemism* appears in the 17th century to refer to words and expressions downplaying the emotional charge of words, which “is characterized by avoidance language and evasive expression” (Allan & Burridge, 1991, p. 3). Euphemisms downplay the emotional impact of words, but also hide or alter their meaning, potentially leading to ambiguity.

Euphemisms can be broken down into different categories, as Jamet and Jobert (2010) point out by enumerating the various existing categorisations of euphemisms based on morphology or semantics. For instance, euphemisms can be distinguished between metaphors, remodellings (like *shoot* for *shit*), circumlocutions (*little girls’ room* for *toilet*), abbreviations (*pee* for *piss*), full-omissions or quasi-omissions, one-for-one substitutions (*bottom* for *arse*), metonymies (*smalls* for *underclothing*), hyperboles (*flight to glory* for *death*), understatements (*sleep* for *die*), borrowing (*anus* for *arsehole*), and technical jargon (*feces* for *shit*) (Allan & Burridge, 1991). However, such categorisations seem both endless and partial, and limit the analysis of euphemisms.

Firstly, the morphological, semantic, stylistic and lexical categorisations tend to draw fixed lists of euphemisms, assuming that some words are always euphemistic. However, as pointed out, social taboos, and their linguistic taboos, evolve over time. A word can be considered euphemistic, but progressively becomes an avoided term, and is replaced by a new euphemism. For instance, the word *undertaker*, once replacing “odd-job man,” was used as an euphemism for people working in funerals, but tends to be now euphemised as *funeral director* (Allan & Burridge, 1991, p. 22). Secondly, a term can be neutral, euphemistic or dysphemistic, depending on its context (Allan & Burridge, 2006; Galli de’ Paratesi, 2009; López Díaz, 2013). For example, the term *operation* is not inherently euphemistic but used in the context of *special military operation* to avoid the term *war*, it becomes a euphemism, and *war* is the taboo word to be avoided. This context-dependence “proves that taboo and euphemism are unstable” (López Díaz, 2013, p. 379). Thirdly, categorisations often disregard the fact that considering a word euphemistic is a matter of evaluation, which Krieg-Planque (2004) terms the “judgment of euphemisation.” It refers to the overlooked value judgement made for terms presented as euphemisms. In dictionaries, some terms are introduced as euphemisms or preferred terms to others, but this type of annotation is not purely linguistic but can be morally evaluative. Fourthly, lists of euphemisms tend to overlook a possible distinction between the

process of euphemising and its results, i.e., the substitute words (Gómez, 2018). In other words, a distinction can be made between “euphemisation” and euphemisms. Fifthly, and drawing from the previous point, euphemisms are not only a lexical process, but also involve grammatical patterns that downplay meaning (Jamet, 2010). Thus, it can be noted that categorisations render partially the complexity of euphemisms. Subsequently, euphemisms are discussed along a continuum of blurriness in this thesis, as explained in the following paragraph.

The continuum of blurriness

Jamet (2010) analyses euphemisms along a degree of blurriness (“floutage”), i.e., the distance between the signifier and the signified. In other words, the blurriness relates to the distance between the term and its referent reality. The greater the distance, the greater the blurriness and, consequently, the stronger the euphemism. As a result, euphemism can change the meaning of the substituted terms by being distant from the direct term, also called *orthophemism* (Allan & Burridge, 2006, p. 33). Thus, euphemisms can be misleading (Allan & Burridge, 1991; López Díaz, 2013; Popescu, 2018a). For instance, by using the term *peace process* to name the war in Iraq, the media choose an antonym to blur the understanding of the events (Haarman & Lombardo, 2011). Klemperer (1947) observes similar linguistic manipulations by the Third Reich in Germany to meet the regime’s ideology. In this case, new taboo words blur reality because common words are censored by the regime. Truth is violated because the word chosen “instead of the forbidden word” does not refer to the same meaning (Chamizo-Domínguez, 2018, p. 80).

However, euphemism can mitigate reality to preserve an individual’s dignity and sensitivity. Widłak (1965) enumerates four reasons for using euphemisms to avoid taboo terms: religious beliefs, decency, compassion, and caution. They all aim to preserve someone’s face as the emotional impact is reduced and euphemised. For example, doctors euphemise diseases and death to preserve their patients’ feelings (Semino et al., 2020; Tayler & Ogden, 2005). On the one hand, this type of euphemism preserves people’s sensitivity and self-images. For instance, “[w]hen members of the black community campaigned to be called African Americans, it was to emphasize not genetics or colour, but the historical roots” (Allan & Burridge, 2006, p. 90). In doing so, euphemisms are motivated by political correctness that encourages “people to change their linguistic behaviour” (Ibid., p. 97). On the other hand, such political correctness may lead to excess. Cameron (1995) reports the anecdote of the automated replacement of *black* by *African American* in a Californian newspaper correcting sentences like “a plan for putting Massachusetts back in the African-American” – instead of keeping the phrase “back in the black” (p. 116). Thus, this kind of “redenomination” motivated by political correctness may be misleading, even though the intentions are meant to be respectful (López Díaz, 2018b). For instance, in French, the words *clochard* (“tramp”) or *pauvre* (“poor”) are socially stigmatised, therefore euphemised into *sans-abri* (“without shelter”) or into the acronym *SDF* (“without fixed domicile”) where the positive implications of “shelter” and “fixed domicile” are simply mitigated through the privative preposition *sans* (“without”) (López Díaz, 2013, p.

384). Therefore, the process of euphemising blurs reality for different reasons, and it may lead to ambiguity.

What is neutral?

Raising the issue of what is considered euphemistic is challenging, as it depends on the speaker's viewpoint. For instance, Cameron (1995) wonders whether the term *collateral damage* is a neutral or euphemistic term in the context of the Iraqi war. If the nominalisation *civilian death* is selected, it would appear as a "euphemistic denial," but if *murder* is chosen, it seems to be biased against the allies. In this case, "it is impossible to come up with a description which could not be interpreted as in some way taking sides" (p. 74). Similarly, when analysing media discourse on incest, it is challenging to identify terms that could be perceived as neutral or euphemistic. Firstly, incest concerns sexuality and, as discussed previously, sexuality is imbued with linguistic taboos. It was noted earlier that neutral terms referring to sexuality, such as scientific terms like *vagina*, can be censored. Thus, it seems that neutrality regarding sexual terms is hardly achievable, and euphemisms prevail in this domain. Secondly, the social taboo of incest gives rise to linguistic taboos and hinders open discussion. Allan and Burridge dedicate a unique paragraph to the taboos of "INCEST and PAEDOPHILIA" in their 1991 book in which they note:

Sex educators and social workers use *touch to children* as a euphemism for *incest*, *pedophilia*, *pederasty*, *interfering with a child*, or *child molestation*. A 1990 poster in New York subways advertising an Incest Hotline reads: 'When someone touches you in a way that feels bad, that's incest.' Next to this is a list of potential touchers: Father, Mother, Uncle, Brother, etc. Euphemistic expressions like *child molestation* or *the sexual abuse of children*, where the nouns derived from the intrinsically dysphemistic verbs *molest* and *abuse* characterize society's abhorrence for the adult perpetrator. (p. 95, emphasis in original)

Thus, they consider the terms *molest* and *abuse* as dysphemisms to *incest* and *paedophilia*. This thesis holds the opposite view: *molest* and *abuse* are euphemisms that obscure the perpetrators' actions and minimise the emotional impact of incest and paedophilia. Chapter 6 outlines the recurrence of the term *abuse* and its inherent vagueness in the context of incest.

2.2. Metalinguistics, vagueness and ambiguity

Metalinguistics

From a metalinguistic perspective, euphemisms are meaningful, and reveal the uneasiness of choosing the right word. Metalinguistics refers to the "discourse about language" (Moschonas & Spitzmüller, 2010, p. 21). Commenting on a euphemism, such as saying in a discourse "it is a euphemism," highlights that a word is used imperfectly or incorrectly to describe the narrated facts (Krieg-Planque, 2004). Such metalinguistic comments are made by journalists to evaluate their own words or reported ones (Ferrarotti, 2011; López Díaz, 2013). For instance, journalists may use a term in inverted commas to distance themselves from it (Quarta, 2009) or question a term through comments such as "let's call it like that" (Ferrarotti, 2011, p. 93). Similarly, in political discourse, politicians may emphasise their anti-

establishment stance in denouncing alleged taboos, regarding immigration for instance (Pražuch, 2018). Comments such as “there is no taboo” or “it is not a taboo” aim to show honesty, responsibility, or even bravery, for openly discussing a topic that is said to be silenced. However, the real impact of such comments is questionable, as the topics are not truly hidden from discussion. On the contrary, the violation of the taboo is highlighted, emphasising how taboo it is. In doing so, “the transgressor, who is not really one, actually does not touch upon any form of the unspeakable” since the topic is openly discussed (Pražuch, 2018, p. 155). Thus, it can be called into question whether metalinguistic comments open the discussion on taboos or reinforce the belief that this topic is too taboo to be discussed.

Regarding incest, the media tend to portray it as a taboo. Years after the abuse, incest victims speak out about the silence imposed by their abusers (Dussy, 2021). Incest perpetrators are aware that incest is socially prohibited, which is why they impose silence on their victims to protect themselves. This silence reinforces sexual violence as “the practice of incest is protected by the absence of words to describe it” (Dussy, 2021, p. 182). Thus, when journalists report victims’ testimonies, they emphasise how the topic of incest touches upon the “unspeakable” (see Chapter 5). Even though journalists praise the courage for “breaking the taboo,” they elude the topic of incest by using metalinguistic comments and avoid analysing the topic in depth. These linguistic avoidance strategies raise the question whether the taboo topic is really reported or only alluded to by emphasising its tabooeness. The taboo may be thus reinforced. However, this caution may be explained by decency to preserve victims’ integrity. Vagueness can play a similar role in language, either to hide or preserve sensitivity.

Vagueness

Vagueness can be defined as “language that is inherently and intentionally imprecise” (Cutting, 2007, p. 4). Vagueness includes not only lexical features but also grammatical ones. Hedging (*like*), approximation (*around*), unspecific words (*stuff*), modals (*may*), “indicate imprecision of word choice” (Cutting, 2007, p. 6). There are different reasons for using vague language. It can create in-group membership by including members who understand the underlying meaning and exclude those who misunderstand it (Cutting, 2007). Vagueness serves purposefully to be unspecific and preserve one’s face (Koester, 2007). Thus, vagueness is context dependent as different meanings may be inferred according to people’s interpretation: “Vague, or indirect, use of language might be beneficial as well in case the speaker is unsure about the preferences of the hearer. In such circumstances a speaker might intent some of his messages to be diversely interpretable” (van Rooij, 2011, p. 5). On the one hand, vague language can be detrimental, such as in trials, when witness report their testimonies imprecisely (Cotterill, 2007). On the other hand, vagueness can be a sign of politeness. As such, students learning a foreign language should learn vague expressions because they are fully part of language (Channell, 1994) and a “language that is overtly explicit can be inappropriate or even rude” (Koester, 2007, p. 58). Therefore, vagueness is not to be understood as either good or bad, but as a linguistic feature that can be “used appropriately” (Channell, 1994, p. 3). In this thesis, vagueness is explored regarding incest. As mentioned above, and later analysed in Chapter 6, the term *abuse* often occurs in the French and British

corpora to name incest. Nevertheless, *abuse* is sometimes too vague to infer the meaning of incest from context. Thus, it is sometimes challenging to ascertain what the journalists meant by using *abuse*. Similarly, Chapter 5 highlights how the term *child abuse* tends to replace both terms *paedophilia* and *incest*. Thus, the vagueness of a term may encompass different meanings and obfuscate the referent reality. It is precisely because terms are vague that they become ambiguous.

Ambiguity

Finally, a distinction can be drawn between vagueness and ambiguity. Vagueness refers to imprecise terms and expressions. Ambiguity refers to the polysemy of a word (Chamizo-Domínguez, 2018), but it may also be syntactic ambiguity. As van Rooij (2011) defines it, “an expression is ambiguous when it has more than one semantically unrelated meaning” (p. 3). For instance, the term *bank* means a finance institution and a river edge. In the sentence “I went to the bank”, *bank* is ambiguous because it could refer to both meanings (Tuggy, 1993). Therefore, ambiguity is defined here as the unclear meaning of a term that the reader cannot determine from the text. In this thesis, two types of ambiguity can be analysed: (i) ambiguity arising from linguistic taboo, i.e., *incest* is sometimes substituted by the vague term *abuse* that may refer to physical, sexual, or emotional abuse (see Chapter 6), (ii) ambiguity arising from social taboo, i.e., incest as intrafamilial sexual abuse is portrayed as being desired. For instance, Chapter 5 analyses how newspapers report eroticised incest, both condemning the taboo and emphasising its attractive transgression. As such, ambiguity arises from a clear ambivalence of how incest is portrayed. The meaning of incest seems to be both negatively condemnable and positively transgressive. These findings are further discussed in Chapters 5 and 6, while the following section expands on the issues regarding taboo and silence.

3. Taboo and Silence

Popescu claims that a linguistic taboo is entirely successful when it leads to silence (Popescu, 2018b). Linguists have taken an increasing interest in silence as they realised that it was an underexplored linguistic phenomenon. Contrary to other disciplines like anthropology or psychology that analyse silence for what it can mean (Tannen & Saville-Troike, 1985), in linguistics, “the tradition has been to define it [silence] negatively—as merely the absence of speech” (Saville-Troike, 1985, p. 3). However, silence should not be simply opposed to noise or speech, as silence can be meaningful in conversations (Jaworski, 1993; Mehrabian, 1981; Tannen & Saville-Troike, 1985), or political discourses (Cameron, 2001; Jaworski, 1993; Sim, 2007). Silence can imply polite disagreement or a desire to avoid naming responsibility (Mehrabian, 1981). Hence, silence cannot be perceived only as a negative behaviour, but it can be positive as well; the meaning of silence depends on pragmatic contexts. For instance, silence can be kept in an argument not to reach a tipping point that the locutor might regret, thus keeping the conversation open for further discussion (Jaworski, 1993). The negative perceptions of silence in Western cultures are “culturally biased” (Ibid., p. 63). Other cultures

value silence, such as in Japan or Nigeria, where it can be perceived as a sign of respect or consent (Tannen & Saville-Troike, 1985).

Therefore, silence should not be bluntly opposed to speech, but placed on the same continuum (Jaworski, 1993). The continuum between silence and speech is noticeable in language itself. Silence is conceptualised as substance through metaphors: “our conventional images of ‘thick,’ ‘deafening,’ ‘heavy,’ or ‘resounding’ silence seem to imply, it often speaks louder than words” (Zerubavel, 2008, p. 8). Thus, silence leaves linguistic marks in discourse, such as metaphors and euphemisms. In addition, Zerubavel (2008) notes “that trauma is only one of the factors that produce silence” (p. 6). This is why silence is worth investigating with regard to taboo language because incest is a trauma that traps its victims into silence (Dussy, 2021). Interestingly, Zerubavel (2008) chooses the figure of Oedipus to illustrate denial when confronted to the violation of a taboo: “upon realising the extent of his own denial, Oedipus actually blinds himself” (p. 4). As such, silence is worth investigating with respect to incest. It should be noted that in the following section, incest is solely defined as intrafamilial child sexual abuse.

3.1. Incest, a conspiracy of silence

In France, half of the incest survivors speak out ten years after the abuse and four out of five families do not protect the child who reveals the incestuous abuse (CIIVISE, 2022). Without exceptions, all the victims of incest report how their abusers ask them to stay silent, underlining the secrecy surrounding incestuous abuse (Brey & Drouar, 2022; CIIVISE, 2022; Dussy, 2021; Gilgun, 1995). In doing so, silence contributes to the domination of incest perpetrators as they make their victims keep silence on the abuse (Armstrong, 1978; Dussy, 2021). This is why it takes many years for victims to finally speak out when they can.

In addition to this domination mechanism, silence enables the reproduction of incest within families. By not believing a victim who speaks out, families deny the frightening reality that a child abuser is among them: “Disparaging the words of incest victims” is “the most common practice of maintaining silence within families” (Dussy, 2021, p. 276). Consequently, silence enables the reproduction of incestuous practices. It has been observed that incest is reproduced over generations in families where incest previously occurred (Brownmiller, 1993; Dussy, 2021; Porter & Ciba Foundation, 1984). Thus, the term “conspiracy of silence” is used to describe the silencing mechanisms surrounding the incest taboo, both within families and within societies (Brey & Drouar, 2022; Dussy, 2021; Justice & Justice, 1980). A conspiracy of silence can be defined as the collaboration between “at least two people” to “avoid acknowledging something” (Zerubavel, 2008, p. 4).

Hence, taboo, denial and silence are closely intertwined. Dussy (2021) explains that families deny and keep silence when they are confronted with the revelation of incest by “‘disbelief’ and ‘respect’ for the head of the family” (p. 180). Similarly, Brownmiller (1993) explains this denial in the larger frame of patriarchy since the “unholy silence that shrouds the interfamily sexual abuse of children and prevents a realistic appraisal of its true incidence and meaning is

rooted in the same patriarchal philosophy of sexual private property” (p. 281). In this regard, the syndrome of parental alienation is one manifestation of this patriarchal domination by not believing the child, and preferring to think that s/he lies and that the hysterical mother seeks revenge against her husband (Brownmiller, 1993; Summers & Campbell, 2022; see Chapter 2). Voices of children are silenced because their subordinate status keeps them in an unheard position, which makes Jaworski (1993) notice that women and children are treated similarly in the ways they are both unseen and unheard. Indeed, women are not believed when they said to be victims of rape (Tranchese, 2023), and children are not believed when they say to be victims of incest. Silence is a sociolinguistic manifestation of oppression. The dominated groups or individuals remain silence as language is dispossessed from them by the dominants (Bourdieu, 2001). Furthermore, when a group is oppressed, “one of the manifestations of their taboo status is that they become silenced,” consequently, “it can be called the *tabooing* of a group to impose silence on it” (Jaworski, 1993, p. 132, emphasis in original).

3.2. Incest, its textual silence

Tellingly, the first book on incest published and promoted on TV in France is entitled *Le Viol du Silence* (“The Rape of Silence,” not translated) by Eva Thomas in 1986. Interestingly, 37 years later, a documentary on incest co-directed by the actress Emmanuelle Béart, who revealed having been an incest victim, was broadcast on French TV and entitled *Un silence si bruyant* (“A such noisy silence,” not translated) (Béart & Mikova, 2023). Here again, *silence* is the keyword. As such, silence relates to the notion of the unspeakable. In a text, silence leads to absence which can be interpreted in different ways. For instance, texts as ancient as the Old Testament show the everlasting triangulation between sex, taboo, and silence. There is one ambiguous passage that is interpreted in different ways by religious and literary scholars because it could depict incest: a scene taking place between Noah and his son Ham in a tent (Rashkow, 2000). Noah is inebriated and undresses himself in front of his son. Ham tells his two other brothers about their father’s nakedness; they come and cover their father. However, when Noah realises it, he is infuriated, and curses Ham. Rashkow comments (2000):

... the severity of Noah’s reaction suggest that the Torah has “*suppressed the sordid details of some repugnant act*” (Sarna 1989:66). But whose “repugnant act”? ... While the narrator is *silent* as to what *actually occurred* between Noah and Ham, the text does report that Ham told his brothers about the encounter. (p. 99, emphases added)

The inability of scholars to agree on the significance of this passage shows that when there is silence, there are ambiguities, and consequently different possible interpretations. Rashkow (2000) considers this passage ambiguous as Noah’s strong reaction either acknowledges “an *actual* incestuous, homosexual encounter with his son Ham,” or “his *repressed* desire for such a relationship” (p. 101, emphasis in original). Thus, the passage is ambiguous and leaves space for interpretation. In addition, this example illustrates the ambivalent feelings towards the incestuous taboo embodied by Noah. Such ambiguity and ambivalence still pervade the ways incest is described in the press, as analysed in Chapters 5 and 6. For this purpose, a

methodology is applied to investigate silence while limiting overinterpretation of what is not present in the texts, as explained in the following Chapter 4.

4. Summary

This chapter demonstrates that sexuality has been submitted to various prohibitions, thus leading to diverse linguistic taboos. Indeed, the social taboo of incest influences the linguistic taboos on incest. Therefore, the notion of taboo language is expanded to include euphemism and vagueness that may lead to ambiguity. As such, this thesis considers taboo language as linguistic avoidance strategies. Furthermore, silence plays a crucial role in understanding incest as an absolute taboo leads to silence. Interestingly, Zerubavel (2008) notes how the taboo topics, underexplored in research, are also “rather difficult to observe” as they “typically consist of nonoccurrences” (p. 13). He adds: “After all, it is much easier to study what people do discuss than what they do not (not to mention the difficulty of telling the difference between simply not talking about something and specifically avoiding it)” (Ibid.). For this reason, the following Chapter 4 explores the methodological challenges and solutions to investigate taboo language, and silence, as its absolute manifestation.

Chapter 4: Research Design

Science does not work by waiting for the perfect data and the perfect method.

(McEnery & Brezina, 2022, p. 64)

In order to investigate the taboo language surrounding incest, as discussed in the previous chapter, the aim of the present chapter is to introduce a novel methodological framework that triangulates both methods and data. As Ehrlich et al. (2014) note, “in terms of language production, and to some degree interpretation, empirical research on the unsaid and unsayable presents a formidable challenge” (p. 208). As such, the methodology addresses the challenges of analysing the unspeakable in news discourse. This chapter first outlines the benefits and limits of combining corpus linguistics and critical discourse analysis, a method triangulation often referred to as corpus-assisted discourse studies (CADS). Within this methodological framework, two underexplored areas are investigated: cross-linguistic CADS and discursive absence. Then, methodological solutions are suggested to investigate cross-linguistically the coverage of incest in the French and British press from 2017 to 2022. Finally, this chapter introduces the corpus of this research. The data has been selected and sorted to pave the way for the following analytical Chapter 5.

1. Corpus Linguistics and Critical Discourse Analysis, a method triangulation

After introducing the benefits and limits of both methods, corpus linguistics and critical discourse analysis, this section explains how combining them, as corpus-assisted discourse studies (CADS), provides a robust method triangulation.

1.1. Corpus Linguistics

Corpus linguistics (henceforth CL) is a computational method to observe linguistic patterns. It aims to analyse “large amounts of linguistic data in a relatively short period of time” (Tranchese, 2023, p. 41). CL is a relatively recent discipline in linguistics, born from computational tools enabling linguists to carry out large-scale analyses on language. It expanded in the 1980s with the advent of personal computers (Baker, 2010, p. 2). Thus, software has been developed in the last decade to collect even larger amounts of data in a shorter time with a wider range of functionalities to study textual data (O’Keeffe & McCarthy, 2022). Among other uses, CL can be used for language description, language teaching, discourse analysis, lexicography, or forensic linguistics (Baker 2006).

However, even though corpus linguistics was born in the 1960s with the first collection of a one-million-word corpus at Brown University in the United States, corpus studies are not new. Corpora have always been analysed, starting with Bible studies where scholars aimed to “point

out the homogeneity of texts (of God) through similar keywords and their frequency” (McCarthy & O’Keeffe, 2010, p. 5). Nowadays, the development of social media like TikTok, WhatsApp, or Instagram, challenges the “text-based paradigm” of corpus linguistics (O’Keeffe & McCarthy, 2022, p. 4) in addressing the inclusion of multimodality to investigate how images, colours and other semiotics interact with written texts (Gillings et al., 2023; Kress & Van Leeuwen, 2001).

The corpus linguistic tools, which are described shortly, provide statistical data on the texts under study. They reveal the occurring linguistic patterns in texts. Two approaches to a corpus can be distinguished. Either researchers want to test hypotheses by looking at the corpus data, or the researchers let themselves be driven by the corpus output. Those two approaches to the corpus are respectively termed “corpus-based” or “corpus-driven” linguistics (Tognini-Bonelli, 2001). However, it has been argued that these approaches are inextricably embedded and should be perceived as the two extremes of the same continuum (Baker, 2010). For instance, researchers may want to check a hypothesis but may be carried in another direction in discovering unexpected data, leading them to new linguistic considerations in their research. However, it is precisely because computational tools are used that such unexpected results can be found. Researchers’ bias is reduced by limiting the cherry-picking of results to fit their theory (Baker, 2006; Baker & Levon, 2015). Another advantage of these tools is that they shed light on linguistic phenomena that would have been more challenging or even impossible to determine through manual analysis due to time-constraints, material limitations, or human capabilities. The four main tools used for corpus linguistic analysis are frequency, keyword, collocation, and concordance. They are briefly listed as they are applied to the analysis in Chapter 5.

Frequency is the “actual count of all occurrences of a particular word in a corpus” (Brezina, 2018, p. 42). Thus, the most frequent terms appear in a frequency list, also called wordlist, indicating whether a pattern in the data is recurrent or rare. For instance, the second most frequent term in the French corpus is *enfant (child)* with 5,384 occurrences. This suggests that the French press tends to focus on children regarding incest. **Keywords** are the most frequent words in the corpus under analysis, relative to a (generally larger) reference corpus. Keywords can be used to determine the aboutness of the texts under study (Baker, 2006), thus, “direct[ing] the researcher to important concepts in a text (in relation to other texts)” (Baker, 2004, p. 347). In the English corpus, the top keyword is *incest*, indicating that this term is more frequent in the corpus under study than in the reference corpus. The details of each corpus are explained in the following section. **Collocation** refers to the “combinations of words that habitually co-occur in texts and corpora” (Brezina, 2018, p. 67). Analysing a word’s collocates can help the researcher to determine the semantic prosody of that word, i.e., the lexical and discursive meaning surrounding it (Gillings et al., 2023; Morley & Partington, 2009). For instance, the term *incest* strongly collocates with *rape (viol)* in both the English and French corpora, suggesting an idea of violence in relation to incest. **Concordance** is another tool that can be used to investigate semantic prosody. Concordance is an interesting tool to “view a search term or phrase within its linguistic environment” (Gillings et al., 2023, p. 21). The search

term or “node term” is displayed at the centre of the screen by the software. The word is observed in context to understand how it is (often) used. A term may seem positive, but its concordance analysis reveals that the node term is always surrounded by negative collocates, like the verb *cause* (Partington et al., 2013; Stubbs, 2001). It is deduced that the verb *cause* is negatively evaluated, thus imbued with negative prosody. Therefore, an appealing benefit of corpus techniques is that they can help to shed light on recurring linguistic patterns that may have been unnoticed without using computational tools (Baker, 2006; Partington et al., 2013).

By now, corpus linguistics has reached “the maturity of the methodology” to reflect on its own limits (Taylor & Marchi, 2018, p. 2). Researchers are therefore called to be aware of these methodological limitations before conducting corpus linguistic analysis. Four main limitations are considered. Firstly, corpus linguistics is often said to conduct too large-scale analysis and to miss information that would enable more fine-grained analysis (Baker, 2006; Baker et al., 2008; Baker & Levon, 2015). Secondly, and related to this, is the impossibility of conducting multimodal discourse analysis with computational tools. Some attempts have been made (Caple et al., 2019), but multimodal analysis often studies small-scale datasets (Bednarek & Caple, 2019; Kress & Van Leeuwen, 2001) because corpus linguistics focuses first and foremost on language (Gillings et al., 2023). In the case of news discourse, it is a limitation not to be able to consider images on a large scale because “newspaper articles are characteristically multimodal and typically contain a mixture of text and visuals” (Brookes & Baker, 2021, p. 257). Thirdly, the usage of software requires a know-how that often involves time-consuming training and financial support (McEnery & Brezina, 2022).

Fourthly, the variations between software can influence the results, and thus the interpretations. Results yielded by different software for the same corpus can vary greatly for two main reasons: the varying definitions of linguistic items, and the diverse calculation methods embedded within the software. For instance, linguistic items like a token are not necessarily defined in the same way across software tools. A token is a word and counting the number of tokens “refers to the total number of words in a text” (Baker, 2010, p. 20). However, should a hyphenated word be counted as one token, or as two tokens? McEnery and Brezina (2022) notice in a survey of tokenisation “that word counts in a corpus can vary up to 17% depending on which corpus tool one uses to count the words in a corpus” (p. 76). Therefore, the calculation measures depend on the selected measurement units. The definition of a token influences the calculation, and consequently, the results. Finding such variations for a basic linguistic item may be even greater for more complex linguistic items, like a collocate. How many words should be around the node word to be considered as a collocate? Which association measure should be selected? The concept of collocation is what Brezina and McEnery (2022) calls an “axiom” in corpus linguistics referring to “what one might call the fundamental assumptions on which a theory is built” (p. 73). An axiom, such as a collocate, is a “provisional truth” (ibid., p. 69) on which the theory of corpus linguistics may develop. Even though the calculation methods depend on how a collocate is defined, those “provisional truths” are accepted by convention to keep investigating language:

Are these methods and are these data the only methods and data that may be used to study a range of linguistic phenomena? No — it is perfectly easy to imagine different sampling structures and even different methods for looking at some of the research questions studied. Does this prove that the sampling structure or the methods in question are ‘wrong’ or ‘deficient’ in some way? If this was done as a prelude to the reach for the perfect set of methods or the perfect data, this would be dangerous and, potentially, the beginning of an intellectual dead end where researchers seek to get data sets and methods ‘right’. (McEnery & Brezina, 2022, p. 64)

Therefore, software like WordSmith, SketchEngine or AntConc, use different calculation methods to analyse the corpus data, thus yielding different results. In order to reduce the variations that may arise from one software tool to another, only Sketch Engine software was used for conducting the whole corpus linguistic analysis of this thesis (see Section 3 of Chapter 4). Thus, the findings are influenced by the measures used by Sketch Engine. Furthermore, what is observable in one corpus is not a general truth, thus some generalisations may be inferred, but always with caution and relativity. McEnery and Brezina (2022) take the example of the white swans that were believed to be only white for centuries in Europe. It was only in 1697 when Dutch explorers saw black swans in Australia that refuted the claim that all swans are white. In the same way as one black swan is enough to contradict a general theory, one contradictory fact in or outside a corpus is enough to undermine generalisation from an analysis. The only general truth that could be inferred is that not all swans are white. Likewise, conclusions drawn from a corpus analysis are provisional because the data extracted from a corpus is the sample of a language that keeps changing and evolving. Thus, McEnery and Brezina recommend considering repeatability “in order to test whether, in the gap between model and reality, there may be found data that confound our hypothesis” (2022, p. 69). To some extent, results are subjected to the researcher’s and the computer’s bias, but the process that leads to these results should apply the principles of replicability and transparency.

1.2. Critical Discourse Analysis

Critical Discourse Analysis (henceforth CDA) broadly aims to shed light on the power at stake in discourse (Fairclough, 2013; Van Dijk, 2001, 2006; Van Leeuwen, 2008; Wodak, 2014; Wodak & Meyer, 2016). CDA draws on Halliday’s Functional Grammar (Halliday & Matthiessen, 2004) to show how sociolinguistic diversities shape social groups and encode their ideologies in discourse (Fowler, 1991). In other words, CDA analyses “choices of grammar and language” to “discover the wider discourses they carry” (Machin & Mayr, 2023, p. 27). CDA primarily defines discourse as social practices and aims to be critical towards who produces discourse, not to investigate the “linguistic unit per se” (Wodak & Meyer, 2016, p. 2). However, discourse is an inherently difficult notion to define because of its evolving and ungrasping nature (Baker, 2023). The notion of discourse generally falls into one of three main definitions: (i) “anything beyond the sentence”, (ii) “language use”, or (iii) “a broader range of social practice” (Tannen et al., 2015, p. 1). However, these three definitions can be challenging to differentiate as they are embedded within one another. Discourse is here defined as the language used in a text shaped by broader social practices. By this, it means that this thesis investigates how the incest

taboo is discussed in the French and British press and contrasts the cross-linguistic findings. As seen in the previous chapter, social taboo generates linguistic taboos that may differ over time and across society, even though some taboos, especially sex taboos, are universal. For this reason, analysing discourse on incest may indicate the sociolinguistic practices of discussing a sexual taboo. Thus, the discourse analysis focuses on (i) the occurring terms through CL tools like frequency and keyword lists, (ii) how the terms are used through collocations and concordances, and (iii) how the terms indicate a broader representation of the incest taboo.

In addition, CDA is “a qualitative framework for the study of language” to analyse small datasets manually (Tranchese, 2023, p. 53), but the framework is politically engaged as well (Fairclough, 2013). The term “critical” should not only be understood as the fact that the discipline aims to critically reveal power influences in discourse, but also as the self-reflection promoted by the discipline itself. Researchers who conduct CDA are also beneficiaries of a position of power (Wodak & Meyer, 2016). Furthermore, CDA is an “interdisciplinary research programme” (Wodak, 2014, p. 302) because it links different disciplines such as sociology, history or philosophy giving “a wider exploration of society and discourse” to critically investigate “linguistic patterns and trends” (Tranchese, 2023, p. 56). This critical stance calls into question whether “a supposedly scientific endeavour should allow itself social and political motivation in the first place” (Chilton, 2005, p. 21). However, the multidisciplinary at the heart of CDA and the commitment to making clear the researcher’s agenda aim to reach greater robustness in analysis (Baker et al., 2013).

Therefore, this thesis is strongly influenced by the political approach of CDA because the topic of child sexual abuse is considered political. The meaning of political should be understood not as politics defended by opposing political parties, but as it involves the whole society’s awareness. As Henry Kempe wrote in 1979, there are six stages on which a society stands when it acknowledges sexual abuse against children. On a continuum of increasing recognition, it goes from complete denial to ensuring care for every child (Porter & Ciba Foundation, 1984, pp. xiii–xiv). The choice of newspapers to raise awareness on child sexual abuse, and especially incest, is determined by how childhood is constructed in society since it is perceived as the epitome of innocence or mischievousness depending on the time periods and economic conditions for children (Jewkes, 2015). Regarding the UK, Jewkes highlights the paradox of “a culture in which romanticized images of childhood as a time of winsome innocence prevail over a reality that includes child abuse, neglect, exploitation and the highest teenage pregnancy rate in Europe” (Ibid., p. 95). Therefore, incest is a political question about how family and children are defined in society.

In addition, the political perception of incest is also determined by external factors such as the “gender dynamics in the news room” (Skidmore, 1998, p. 213), i.e., the prevalence of male journalists in news production. CDA sees discourse as “sites of struggle” for dominance (Wodak & Meyer, 2016, p. 10), and news discourse is not the product of one individual, but of different stakeholders including reporters, journalists, editors, and investors (Fowler, 1991;

Richardson, 2007). CDA investigates the political agenda of news production, because news are not neutrally transmitted, but are selected according to the predominant ideologies of newspapers (Dijk, 1988; Fowler, 1991). Newspapers are an interesting field of investigation for practical reasons, as a large amount of data is rather easily accessible and collectable, and for sociolinguistic reasons as news discourse influences and is influenced by society (Bednarek & Caple, 2019). News discourse refers to the language in the mass media that construes news (Bednarek & Caple, 2019). Events are selected and transformed by the media to become newsworthy. Events happening in societies are not “intrinsically newsworthy, but only become ‘news’ when selected for inclusion in news reports” (Fowler, 1991, p. 11). Newsworthiness refers to the news values that “determine what makes something newsworthy – worthy of being news” (Bednarek & Caple, 2019, p. 50). News is foregrounded or backgrounded (or completely absent) depending on their news value. Bednarek and Caple (2019) enumerate nine news values: negativity, timeliness, proximity, prominence, consonance, impact, novelty, superlativeness, personalisation. These values are commented on in the following analytical chapters as newsworthiness is worth being investigated in the framework of CDA to understand how incest is, or is not, covered in the press.

Research conducted in the late 1990s in the UK is worth mentioning to understand the social construction of news regarding incest, and child sexual abuse in general. Carter et al. (1998) show that the media coverage of child sexual abuse is avoided because of male journalists’ reluctance to cover such events. Some journalists explain it by the fear of “child abuse fatigue” referring both to their own boredom and the readership’s boredom of regularly reading stories on child abuse. An interviewed journalist claims that “everyone was getting sick with the child-abuse story” even though such evidence is not provided (Ibid., p. 194). Furthermore, in a male-dominated newsroom at that time, the rare female journalists struggle between the fear of being confined in news relating to families and children and the interest for covering that type of news with credibility. Whenever female journalists believed that an alleged culprit was an actual abuser, their feelings “were dismissed by male editors as ‘subjective’ or ‘biased’ whilst their male colleague’s ‘gut feelings’ in favour of the same man, were seen to constitute ‘common sense’ or ‘professional instinct’” (Kitzinger, 1998, p. 198). To this gender bias, it should be added a liberal theory that promotes the myth of the families and blames the failure of healthcare system to prevent child sexual abuse – and where the staff is predominantly female: “the liberal, broadsheet thesis deflected attention from masculinity and family towards bureaucratic failure” (Wykes, 1998, p. 244). The sacredness of house is preserved since families are never described as “a site of crisis for many battered women or abused children [...] yet home and family are still promoted as the desirable norm” (Ibid., p. 235). Furthermore, common abuse reported in police reports is underrepresented in the press, but the rarer events such as feminicides are overrepresented, which makes Carter (1998) say that the “extraordinary” becomes the “ordinary” in everyday news. In her research, Carter observes that in a sample of articles on sexual violence, 3% covered incest, while 42% feminicides, yet she argues that incest is more common than feminicides, suggesting that the press normalises the “extraordinary” rather than the “ordinary.” The mismatch between the police reports and the press shows that “a desire for the ‘unusual’ had led to increasingly

sensational forms of sex crime reporting” (Ibid., p. 229). Similarly, Brownmiller (1993) notes “that sexual offenses against children are barely noticed except in the most violent and sensational instances. Most offenses are never revealed, and when revealed, most are either ignored or not reported” (p. 277). That is the reason why the notion of newsworthiness is worth discussing further, especially in the context of CDA.

Bednarek and Caple (2014) notice that news values are not enough quoted in CDA studies. They consider that the analytical concepts such as social actors, transitivity, or modality, are constantly used, whereas the notion of newsworthiness is barely mentioned in various studies on news media. They explain it by a lack of interdisciplinarity between journalistic studies and linguistic studies, but also by a “lack of an appropriate systematic analytical framework for the linguistic analysis of news values” (Ibid., p. 139), which they address in their article. Analysing news values gives “insight into how happenings are ‘sold’ to us as newsworthy through conventionalised ways of saying that are repeated frequently” (Ibid. 2014, p. 143). Notably, it has been observed that children become newsworthy when they are related to sex crime (Greer, 2003; Jewkes, 2015). Therefore, this thesis aims to address news values through the comparison of the French and British press to understand whether incest is newsworthy depending on the political agenda. For this reason, corpus linguistics and critical discourse analysis are triangulated to analyse media coverage of incest.

1.3. Method Triangulation

Triangulation “is a term taken from land surveying which uses distance from and direction to two landmarks in order to elicit bearings on the location of a third point (hence completing one triangle)” (Baker & Egbert, 2016, p. 3). Taylor and Marchi (2018) distinguish four types of triangulation in linguistics: methods, tools, investigators, and data. This thesis uses two types of triangulation combining two different methods and two sets of data. This section focuses on the method triangulation between Critical Discourse Analysis and Corpus Linguistics. Data triangulation is explained in the next section.

Method triangulation between CDA and CL has been largely documented and applied (Baker, 2006; Baker et al., 2013; Brookes & Baker, 2021; Gillings et al., 2023; Hardt-Mautner, 1995; Partington et al., 2013; Taylor, 2014; Tranchese, 2023). Hardt-Mautner (1995) was the first to use a combination of corpus linguistics and critical discourse analysis for conducting a media analysis. She noticed that the amount of data cannot be analysed through CDA alone; it had to be complemented with a more quantitative approach to cope with it. Her analysis of News Actors was facilitated by combining CDA with CL tools. Her article illustrates an early age in the combination of CDA with CL, but today, this combination is extensively developed and has taken a step further. Baker (2006, 2023) mainly contributed to systematising the approaches of corpus linguistic with discourse analysis.

Combining CL and CDA has been argued to mutually counterbalance each method’s respective limitations and reinforce their strengths. CL yields sound statistical results, while CDA offers a greater social and political context to the findings (Brown, 2023). CL provides computational

tools for building a corpus and observing salient, sometimes unexpected, linguistic patterns, while CDA tend to “cherry-pick” its data and avoid taking into account occurrences that contradict the researcher’s hypotheses (Baker, 2023; Baker et al., 2008; Baker & Levon, 2015; Gillings et al., 2023). Thus, the main advantage of the method of triangulation is to reduce the researcher’s bias. Additionally, results from two methodologies applied to the same dataset may either converge or complement each other, confirming the hypothesis; but in the case of contradictory results, investigating the dissonance is necessary to understand its cause (Marchi & Taylor, 2016). Thus, method triangulation provides greater validity of the results (Baker & Egbert, 2016; Marchi & Taylor, 2016; Tranchese, 2023). Moreover, replicability is facilitated as researchers “make their choice of corpora and analytical techniques transparent” (Brezina, 2018, p. 2).

As a result, the methodology of corpus-assisted discourse studies (CADS) was developed through the combination of CL and CDA. Even though Partington (2013) coined the term CADS and defined this research method as “not tied to any particular school of discourse analysis, certainly not, for instance to *critical* discourse analysis” (p. 10, emphasis in original), other scholars use the term CADS for referring to the combination of CL and CDA specifically (Baker, 2023; Baker et al., 2008, 2013, 2021; Brookes & Baker, 2021; Tranchese, 2023; Vessey, 2013). As such, CADS can conduct both “critical and non-critical types of research” (Baker et al., 2021, p. 53). However, the goal remains to be critically engaged with the data. For instance, a corpus-assisted discourse analysis can “be considered in light of broader social issues relating to news media production and sexual violence” (Tranchese, 2023, p. 57). In this regard, this thesis aims to address the sociopolitical meaning of the media discourse on the incest taboo. For these reasons, this thesis conducts a corpus linguistic analysis informed by critical discourse analysis.

However, this “useful methodological synergy” of CADS (Baker et al., 2008) relies on the “key assumption” that CL is a purely quantitative method whereas CDA is qualitative, leading to a “misapprehension” of both methods (Taylor & Marchi, 2018, p. 2). This “misapprehension” is due to a lack of understanding of what both methodologies allow researchers to do. CL relies on quantitative findings, but also on qualitative interpretation, because data is selected and interpreted (Baker, 2006; Taylor & Marchi, 2018). As Marchi and Taylor (2016) note, “a corpus approach is not a neutral one, selection comes into play throughout the process and early stages of analysis (e.g. the choice of the lexical items to investigate) can heavily determine the progress of the research” (p. 4). Even though CL software reduces the researcher’s influence, there are still decisions to make that require human decisions. For instance, cut-off points are a critical measure in corpus linguistics, determining the threshold at which a word is revealing (Baker, 2023), whether it is set at 10, 100, 137. These cut-off points tend to be round numbers, influenced by cognitive bias that may affect the analysis. Furthermore, the analysis of the data extracted through CL tools is interpreted by the researcher. Therefore, rather than opposing CL and CDA, they should be apprehended on a “continuum” (Taylor & Marchi, 2018, p. 6).

Another assumption to debate is that combining two methodologies guarantees greater objectivity (Taylor & Marchi, 2018). However, combining methods cannot guarantee

neutrality, as the choice of topic is often motivated by the researcher's personal interests, implying subjectivity from the very beginning (Marchi & Taylor, 2016, p. 4). Furthermore, there is no guarantee of reliability since the researchers may combine flaws of each discipline (Ibid.). Indeed, two methods, instead of one, may ignore facts that contradict the hypotheses. As such, cognitive bias such as confirmation bias may be reinforced by the combination of two methods. Finally, the way the researcher interprets the results is not necessarily replicable, even though the method used is (Taylor & Marchi, 2018). Therefore, the method triangulation of CADS is combined with another triangulation in this research: the data triangulation. The data is contrasted between the English and French languages. Cross-linguistic analysis raises methodological challenges in CADS, as discussed in the following section.

2. Methodological challenges

This section addresses the methodological challenges of conducting cross-linguistic CADS. The aim of this double triangulation, between method and data triangulation, is ultimately to tackle the research question of analysing taboo language, including textual silence. Thus, a methodological framework is introduced to shed light on two overlooked areas in CADS.

2.1. The challenges of cross-linguistic CADS

Cross-linguistic CADS is an underexplored area (Freake, 2012; Freake et al., 2011; Nardone, 2016; Taylor, 2014; Vessey, 2013). However, the "lack of means of direct comparison" should not deter researchers; instead, it should motivate them to be creative in developing analytical tools (Vessey, 2013). This section relies on the work of Vessey (2013) and Taylor (2014) who highlight the methodological challenges of conducting cross-linguistic CADS. Their work raises five methodological challenges, summarised in Figure 4.1 at the end of this section.

The first challenge of cross-linguistic CADS is to build comparable corpora (Taylor, 2014; Vessey, 2013). The corpus built in one language may not be exactly comparable to the corpus built in another language with the same search terms as the terms may seem to be equivalent, but their semantic prosodies differ. Semantic prosody refers to the "evaluative value of the terms" (Taylor, 2014, p. 374) and the "principle of co-selection" of words that keep a discourse together (Morley & Partington, 2009, p. 139). For instance, the terms *incest* in English and *inceste* in French seem to be identical, but the semantic prosody differs between the two languages. Indeed, the findings in Chapter 5 suggest that *inceste* in French primarily refers to intrafamilial child sexual abuse while *incest* in English primarily implies consensual intrafamilial sexual relationships. As such, the terms *incest* and *inceste* occur with different collocations and are embedded in different linguistic patterns, showing distinct sociolinguistic features. For this reason, semantic prosody is culturally determined and it cannot be assumed that the translation of the terms makes them equivalent (Grabowski, 2022; Vessey, 2013). This study focuses on semantic prosody in Chapter 5, as it is the key to understanding the cross-cultural perceptions of incest between French and English languages. One solution can be to select the same text-type to contrast the same term in both languages in similar contexts. In

this regard, this thesis selected daily newspapers in both France and the UK as they belong to the same genre, undergo similar production processes and serve the same purpose.

The second challenge is the risk of essentialising and reifying categories (Vessey, 2013). In this thesis, categories would refer to cultures. What are the French and British cultures that French and English languages are supposed to represent? Language construes the idea of nationhood (Anderson, 2016; Freake et al., 2011). However, various social groups speak the same language (e.g. North Americans vs. Britons, French Canadians vs. French), and inferring generalities is not considering the diversity of the population and its culture(s). Culture is an individual and social construct (Spencer-Oatey, 2012). Defining culture is fuzzy because it is difficult to determine “absolute set of features” that would be seen as prototypical of a cultural behaviour or habit (Ibid., p. 9). Furthermore, essentialising runs the risk of being oversimplistic by overlooking subtleties and differences (Vessey, 2013). The additional risk of essentialisation is focusing on cross-cultural differences rather than similarities, which are often overlooked in corpus linguistics (Baker, 2010; Schröter & Taylor, 2018; Taylor, 2013). However, categorising is a useful tool for conducting an analysis, provided that it is complemented by other approaches (Baker, 2010). Therefore, the term *culture* is used in this thesis in the sense of *nation* as an entry point to comprehend how the French and English languages, used in the press that may shape and mirror their respective society. Indeed, newspapers are one of the “technical means ‘re-presenting’ the kind of imagined community that is the nation” (Anderson, 2016, p. 25). Thus, what may be inferred from the results is what the media *represents culturally, not what culture is*.

The third challenge is cross-cultural variation, i.e., the fact that some “cultural elements are not transferable from one language to another” (Taylor, 2014, p. 374). The main cross-cultural variation to consider in this study is the absence of tabloids in the French media landscape. There is a general distinction between tabloids and broadsheets in the British press. Whereas the latter are perceived as serious journalism or “elite newspapers” (Van Den Bulck et al., 2017, p. 51), tabloids “produce news items that are accessible for everyone” (O’Neill, 2012, p. 27), and may be referred to as “popular newspapers” (Van Den Bulck et al., 2017, p. 52). Unlike the UK, France does not have tabloids. Instead, a differentiation is made between national newspapers that cover both national and international news, and regional newspapers that focus on national and local news. It is worth noting that regional titles sell more newspapers than national titles. For instance, the most widely distributed regional daily, *Ouest-France*, sold 610,637 copies in the year 2023, whereas the most widely distributed national daily is *Le Monde*, which sold 482,756 copies in the same year.² Kuhn (2011) suggests that this “low circulation nationwide” of French daily newspapers is due to “the absence of a popular title with tabloid journalistic values: France notably lacks the equivalent of *The Sun* in Britain or *Bild* in Germany” (p. 35). The success of the regional newspapers can be attributed to their

² The French agency Alliance for the press and media numbers (ACPM) measures and publishes audience ratings on press, radios, television, podcasts. Its website is available at <https://www.acpm.fr>.

coverage of local events, appealing to readership in proximity. In contrast, national titles are produced in Paris and their main readership is in Paris (Ibid.). Therefore, it should be borne in mind that differences in the style of media coverage might be explained not only by language but also by the media context. In other words, the variations from one language to another might be explained by linguistic reasons, but also stylistic reasons. News values attributed to tabloids such as celebrity and proximity may influence a linguistic style that is perhaps not reproduced in French newspapers because of the absence of tabloids. As Vessey (2013) notes, we should “be careful not to assume that each language corpus is a homogenous entity” (p. 375). Thus, linguistic variations may be influenced by broader cross-cultural practices, which are a focus of interest.

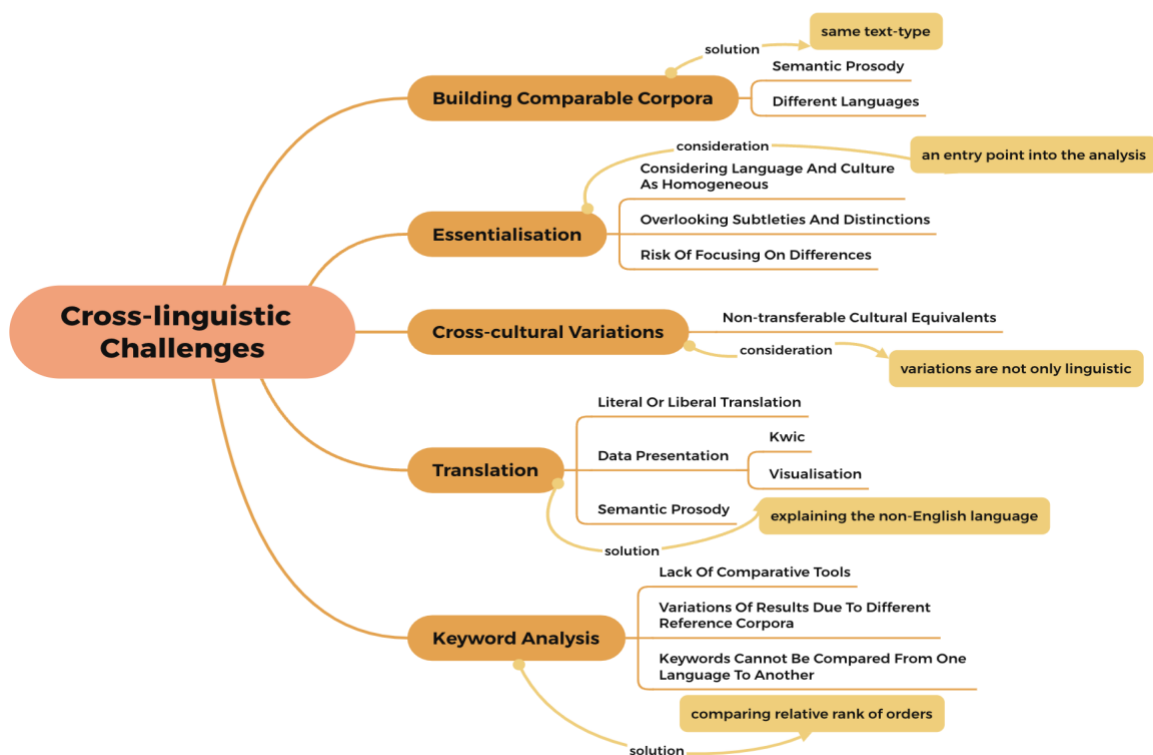
The fourth challenge pertains to translation (Taylor, 2014; Vessey, 2013). There are three main considerations to bear in mind when it comes to translation. Firstly, it raises the question whether data should be literally or liberally translated (Taylor, 2014). Taylor takes the example of the Italian word *strisciante* that could be either literally translated into English as *creeping* or liberally translated as *underlying*. The literal translation “would be less easily understood by the English-speaking audience,” but the functional translation would alter the metaphorical use of this collocate of *racism* (Ibid., p. 374). Consequently, the second question raised is how data should be presented. KWIC (Keyword In Context) and concordances display the node term in the middle of the screen and highlight repeated patterns, but the data may be not understood by readers who do not speak the language. However, translating concordances to make them accessible may change the word order, thus tampering results. It has been observed that the usual practice in cross-linguistic discourse analysis is to present data in the original language and provide translation in the body of text with an explanation (Freake, 2012; Freake et al., 2011; Jaworska, 2014, 2021; Jaworska & Krishnamurthy, 2012; Jaworska & Leuschner, 2018; Taboada, 2013; Taylor, 2014; Vessey, 2013). Hence, this thesis presents data in French with English translations in the body of text. It is worth noting that some visualisations from Sketch Engine are presented as extracted directly from the software, as it cannot translate the French data into English in diagrams, tables and figures. Finally, the third question touches upon semantic prosody, as mentioned above. The translation of a word imperfectly captures its meaning from one language to another (Grabowski, 2022; Taylor, 2014; Vessey, 2013). For this reason, the analysis of the semantic prosody in Chapter 5 reveals the deeper meaning of the comparable terms.

Finally, the fifth challenge is conducting keyword analysis. Generally, there is an “imbalance in developments across English and French corpus linguistics” (Freake et al., 2011, p. 43), and more specifically, there is a lack of comparative tools, such as comparable reference corpora (Vessey, 2013). As briefly mentioned above, a reference corpus refers to a larger corpus used to compare another corpus in order to identify keywords that indicate its topicality (Baker, 2006; Taylor & Marchi, 2018; Tranchese, 2023). Comparing the same corpus to different reference corpora yields different keywords (Baker, 2004, 2023). It means that comparing a French corpus to a French reference corpus does not produce the same results as an English corpus compared to an English reference corpus. For this study, the software Sketch Engine

was chosen over software such as AntConc or LancsBox because it offers two rather parallel reference corpora with its TenTen Corpus Family (Jakubíček et al., 2013). TenTen Corpus Family provides updated and large-scale data with an emphasis on the media. Data was scraped from the internet while avoiding duplicates. The English Web Corpus 2021 (henceforth EnTenTen21) and the French Web Corpus 2023 (henceforth FrTenTen23) – the latest versions available – amount to 52 billion words and 23.8 billion words respectively. As can be noticed, despite both reference corpora undergoing a similar building process, FrTenTen23 is two years more up to date than EnTenTen21, but it is twice as small. Therefore, cross-linguistic keyword analysis is to be considered with caution. In addition, keywords from each corpus cannot be directly compared to each other because the languages are different (Vessey, 2013) and the software is not able to tag them cross-linguistically. A possible solution is to conduct a “parallel keyword analysis” (Ibid., p. 16) which consists in looking at each list and comparing word rankings. In this thesis, another methodological solution is to focus more on the analysis of the wordlist than the list of keywords to avoid the variations of external data, as further explained in Chapter 5.

As can be noted, these five methodological challenges can be addressed and partly solved to conduct a cross-linguistic and cross-cultural analysis. Furthermore, Bednarek and Caple (2017) encourage conducting cross-cultural comparison as an extension of their analysis of English-language news values to raise questions such as “What news values are discursively construed in different cultures? How is the same topic or event constructed as newsworthy?” (p. 8). Those questions are addressed in this thesis.

Figure 4.1 The five challenges of cross-linguistic CADS, adapted from Vessey (2013) and Taylor (2014)



2.2. Finding absence: a “dusty corner” of CADS

Another methodological challenge in the analysis of the incest taboo is exploring absence. Absence is a “dusty-corner” in CADS in the sense that it is an overlooked topic (Taylor & Marchi, 2018). One reason is the fact that CADS is “primed to look at what is *there*” (Gillings et al., 2023, p. 45, emphasis in original), i.e., CADS “tends to focus on what *has* been explicitly written, rather than what *could have been* written” (Partington, 2014, p. 120, emphasis in original). As outlined above, CL tools, like collocations and concordances, are designed to find linguistic patterns in a corpus. At first sight, they do not seem suitable for finding absence, as perhaps CDA does through the analysis of missing agency for example (Partington, 2014; Van Leeuwen, 1996).

Another reason for it being a dusty corner is the lack of empirical studies on silence and absence, as addressed in Schröter & Taylor (2018). Their study provides various approaches for analysing absence in discourse and makes a distinction between absence and silence. Absence is an umbrella term that encompasses all forms of absence in discourse, including silence, as well as a “process of choosing” what is foregrounded or backgrounded (Ibid., p. 7). In this sense, absence and presence are two sides of the same coin where one element is chosen over another. For instance, news values may explain absence and presence in the news discourse. In contrast, silence refers to “a more conscious and intentional choice about what (not) to say” (Schröter & Taylor, 2018, p. 7). However, silence should not be perceived only negatively, since silence is meaningful, and sometimes even respectful (Jaworski, 1993, see also Chapter 2 in this thesis for further discussion on silence). In the case of incest, silence can also protect victims’ dignity.

Drawing from these definitions of silence and absence, this thesis introduces its own definitions. Silence is defined as a coercive process of not saying. For instance, silence is imposed on the victims by their abusers and the families may be in denial of acknowledging incestuous abuse (see Chapter 3). As such, silence is an intentional choice of not saying. Consequently, silence generates discursive absence. Thus, discursive absence is defined as the consequence of silence. A topic is silenced due to its taboo nature; thus, it is absent in discourse. Echoing the words of Schröter and Taylor (2018), discursive absence is a process of choosing what is backgrounded. Furthermore, a distinction can be made between *known absence* and *unknown absence*. Partington (2014) defines known absence as the expected absent features that the researcher aims to investigate, while *unknown absence* refers to an unexpected discovery of absent features that consequently become researchable. For instance, the absence of recognition of the Arab Spring in the UK and US press and political discourses corresponds to his definition of *known absence*, or suspected/searchable absence, that is, “you already know which linguistic feature you are searching for and simply want to know whether or not it is in the corpus” (Partington, 2018, p. 39). In other words, *known absence* depends on a linguistic feature or a political issue that the researcher is aware of and is searching for. Similarly, this thesis investigates whether the incest taboo is portrayed as

intrafamilial child abuse that is often silenced. Thus, from the assumption that this topic is silenced, the analysis examines a *known absence* in discourse.

Another reason why absence is overlooked in CADS is that identifying absence in discourse is paradoxical “because, by definition, it cannot really be said to be anywhere” but “it is everywhere” (Duguid & Partington, 2018, p. 39). It is worth noting that such a paradox has also been raised for taboos that have “noticeable presence” due to their absence (Rosewarne, 2013, p. viii, see chapter 2). In other words, the paradox lies in the fact that anything that is not said is significantly absent. However, the risk is overinterpreting absence because there is “the temptation to over-read, over-claim and over-dramatise when the evidence is limited” (Partington, 2014, p. 135). Therefore, methodological questions must be posed: “How do we come to notice absences? How can we argue the existence of absences, what shapes they take, where and why they occur? What sense can we make of them, how do they determine what is present?” (Schröter & Taylor, 2018, p. 1).

To address those questions, methods in CL have been developed to compare corpora to determine whether certain elements are missing from one dataset to another. Contrasting corpora is at the heart of CL since it aims to investigate linguistic patterns that are otherwise kept hidden (Partington, 2014). Comparing datasets may highlight what is relatively absent from one corpus to another (Duguid & Partington, 2018; Partington, 2014). In other words, “alternative presence” is used to point out absence, otherwise, absence may be overinterpreted or stay unnoticed (Schröter & Taylor, 2018, p. 6). For instance, the method of diachronic CADS compares corpora to determine the historical lexical change in discourse, and consequently, the relative absence and presence of words over time (McEnery & Baker, 2017; Partington, 2018).

Cross-media comparisons are an interesting method to contrast different datasets to investigate what is relatively present or absent. For instance, Alcántara-Plá and Ruiz-Sánchez (2018) notice that the topic of immigration is absent from the twitter accounts of the Spanish politicians running for the general election, while the topic was making headlines in the mass media. They interpret this discrepancy as a way of silencing the immigration topic to garner votes. Conversely, Strand (2018) notices the absence and passivation of the LGBTQ+ community in state-owned newspapers in Uganda compared to their active voices expressed on Twitter. Such cross-media comparison “contributes to the identification of silencing practices and discursive absences” (Ibid., p. 127) while safeguarding against overinterpretation.

Another methodological solution to limit the risk of overinterpretation is to be aware of the social, cultural, historical and political contexts “since exclusion and subsequent absence from discourse are always context-dependent, an understanding of context is pivotal” (Strand, 2018, p. 127). It is worth noting that being aware of the socio-cultural background is one of the strengths of CDA. For instance, Taylor (2014) investigates the “real-world data” of immigration and compares it with the representation of migration in the press to determine if there is any “mismatch” between those two datasets.

Therefore, it should be noted that those different approaches for finding discursive absence rely on the choice of comparable corpora and the contrastive corpus techniques shed light on what is absence (Duguid & Partington, 2018, p. 56). Interestingly, no studies have attempted to suggest a cross-linguistic corpus-assisted discourse analysis to investigate discursive absence. Therefore, this research aims to fill this gap by introducing methodological innovation, which is further explained in the following section.

2.3. Data triangulation, a possible methodological solution

As discussed above, triangulation is a methodological approach to provide greater scientific robustness to research. In this thesis, method triangulation is associated with data triangulation to investigate the linguistic taboos of incest, and by extension, whether the incest taboo is discursively absent. Therefore, this study aims to address the methodological challenges in CADS mentioned above by using two types of triangulations. This double triangulation is summarised in Figure 4.2.

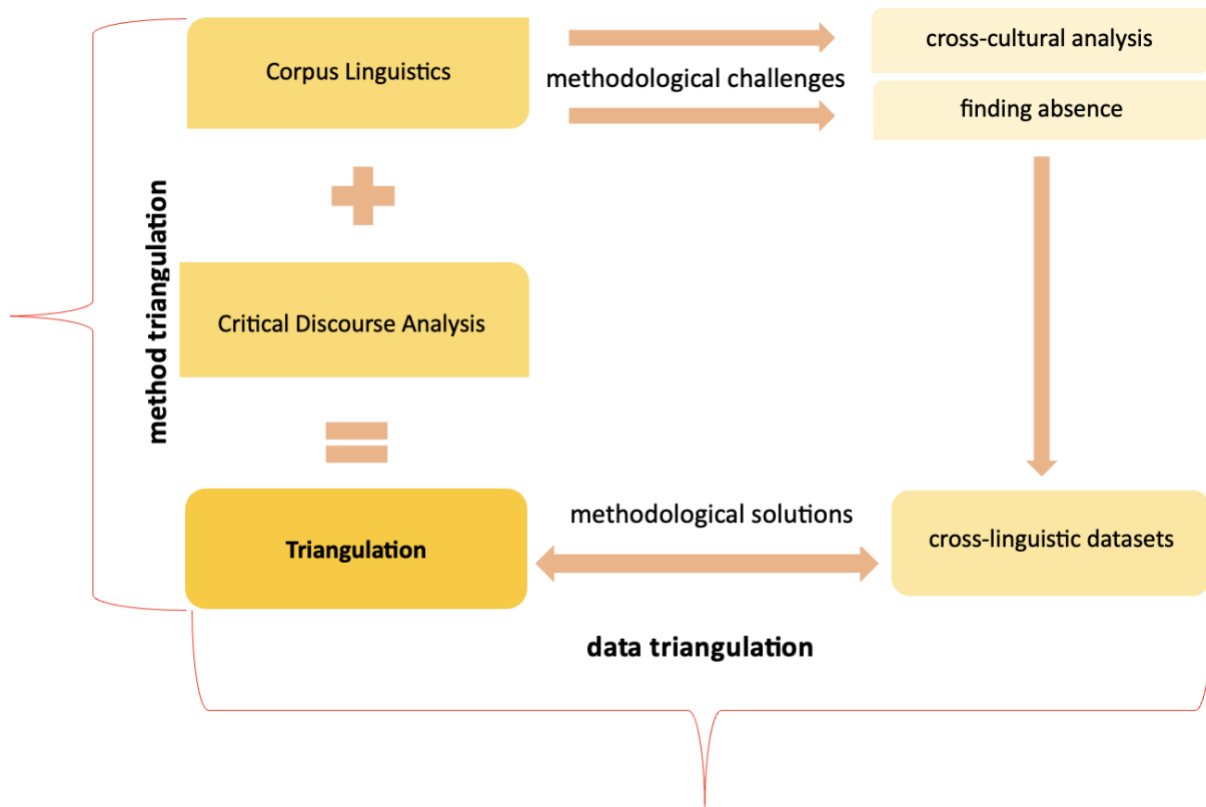
The method triangulation between CL and CDA is an additional approach to cautiously analyse discursive absence because CDA alone often “[takes] for granted that a silence or absent message or voice must have been deliberately *suppressed* with little evidence of intentionality” (Duguid & Partington, 2018, p. 39, emphasis in original). Therefore, revealing absence may be motivated by political reasons when researchers argue they “have some special insight due to their particular politicised worldview” (Partington, 2014, p. 134). As discussed above, CDA has an emancipatory agenda that “seeks to highlight structures within societies that can result in abuses of power and to consider the perspectives and voices of those who are often not regarded” (Brookes & Baker, 2021, p. 7). However, if the socio-political context is not considered, “how can we analytically engage with something that may not be linguistically present and argue the existence of the non-existent voice?” (Strand, 2018, p. 126). Thus, discourse is shaped by social practices that are worth being analysed (Schröter & Taylor, 2018).

As such, newspapers are a place of struggle, and what is covered in the media, and whose voices are portrayed, is a result of those powers at stake. Social and political factors shape news media, and events are considered newsworthy when they align with what society might deem worthy of interest. This thesis demonstrates that the focus on incest in the press is shaped by such social practices. The change of perspective after a high-profile scandal in France (see Chapter 5) shows that “unless social contexts change, it can be very difficult to resist the suppression of alternative versions of social reality, of the viewpoints of marginalised groups, of tabooed narratives” (Schröter & Taylor, 2018, p. 4). Furthermore, the focus on news values aims to reveal what is foregrounded or backgrounded by the press in each country.

Furthermore, CL provides valuable tools to be combined with CDA to analyse absence. Tools such as keywords, collocations, and semantic prosodies, are useful indicators for revealing expected and unexpected linguistic patterns. Therefore, cross-linguistic CADS can (i) identify

absence, i.e., knowing what is absent from the socio-cultural context and other extratextual data, (ii) locate absence, i.e., knowing where it should have been, and (iii) quantify absence, i.e., statistical elements from the computational tools (Partington, 2014). Therefore, CADs offer an answer to the “what, where, how much and the why of absence” (Ibid., p. 142). This thesis aims to answer this question related to silence on the incest taboo in conducting a cross-linguistic corpus-assisted discourse analysis. The corpus used to this end is introduced below.

Figure 4.2 Methodological framework consisting in combining two types of triangulation



3. Corpus Building

3.1. Data collection

The present study examines the representation of incest in British and French newspapers from October 15, 2017, to October 14, 2022, covering the emergence of the #MeToo movement up to the date of the first data collection day. The corpus was built through two search queries, resulting in two corpora in each language. The first search query was *incest** for building comparable corpora in French and English. The second search query, “*abuse AND father*” and “*abus AND père,*” was used to build another comparable corpus in both languages. The second corpus aims to highlight the linguistic patterns of absence on incest. Further details of this corpus building are provided in Chapter 6. In this section, the focus is on the corpus *incest** built in two different languages and two different media landscapes.

The newspaper selection differed between the UK and France due to cross-cultural variations in the media landscape, as discussed above. Indeed, there are no tabloids in France, unlike in

the UK. Instead, a differentiation is made between national newspapers and regional newspapers. Thus, eight British newspapers were selected based on their circulation figures (Media Reform Coalition, 2021; Office of Communications, 2019) and accessibility on LexisNexis.³ Thirteen French newspapers were selected according to their circulation and regional representation of mainland France, as well as their availability on LexisNexis.⁴ Thus, the English corpus includes four broadsheets and four tabloids (Table 4.1). The French corpus includes five national newspapers and eight regional newspapers (Table 4.2).

Table 4.1 Selected British newspapers

| Broadsheets | Tabloids |
|----------------------------|---------------------|
| <i>The Daily Telegraph</i> | <i>Daily Mail</i> |
| <i>The Guardian</i> | <i>Daily Mirror</i> |
| <i>The Independent</i> | <i>The Express</i> |
| <i>The Times</i> | <i>The Sun</i> |

Table 4.2 Selected French newspapers

| National newspapers | Regional newspapers |
|---------------------|---------------------------|
| <i>La Croix</i> | <i>Le Dauphiné libéré</i> |
| <i>Les Échos</i> | <i>La Dépêche du Midi</i> |
| <i>Le Figaro</i> | <i>L'Est républicain</i> |
| <i>L'Humanité</i> | <i>La Montagne</i> |
| <i>Libération</i> | <i>Ouest-France</i> |
| | <i>Le Progrès</i> |
| | <i>Sud-Ouest</i> |
| | <i>La Voix du Nord</i> |

A recurring methodological question in corpus linguistics is to wonder how large a corpus should be. Some may consider that the larger, the better, while others find it more difficult to conduct manual analysis on a large amount of data (McEnery & Brookes, 2022). Nevertheless, a common answer is that the “corpus size depends on the research question and the kind of linguistic features we want to investigate” (Brezina, 2018, p. 18). For this reason, the corpus is built bearing in mind three points to contrast the French and British press to address the research question:

- (i) *Finding a common starting point.* The date 13 October 2017 is selected as the first data collection day. It corresponds to the date of the tweet ending with the hashtag #MeToo from the actress Alyssa Milano following the publication of articles in the

³ LexisNexis Database is available at <https://www.lexisnexis.co.uk>.

⁴ It should be noted that the main French daily *Le Monde* was not included due to its inaccessibility on the database.

New York Times (Kantor & Twohey, 2017) and the *New Yorker* (R. Farrow, 2017) reporting the systematic sexual harassment and abuse by the American producer Harvey Weinstein against actresses in Hollywood. The #MeToo movement started in the United States but found massive echoes in many countries, among which France and the UK (De Benedictis et al., 2019; Gill & Orgad, 2018; Tranchese, 2023). #MeToo was therefore considered as a relevant starting point of comparison between the two countries. The movement is further discussed in Chapter 7 with the case study dedicated to the incestuous sexual abuse allegations by Dylan Farrow, Woody's Allen adoptive daughter.

- (ii) *Aiming for a representative corpus.* Representativeness is the aim of a balanced corpus (McEnery & Brookes, 2022), even though a corpus is inherently imperfect as language keeps evolving and what is captured at one moment has already changed the moment after (Brezina, 2018). However, balance is a goal to attempt to guard against distortion (Gillings et al., 2023). For this reason, only daily newspapers were selected in both countries. Even though tabloids do not exist in the French media landscape, the dichotomy between regional and national newspapers is considered. Furthermore, the political leanings of the newspapers are not taken into account because they differ between the UK and France. For instance, there are more left-wing newspapers in France than in the UK, especially with the two strongly engaged left-wing newspapers *L'Humanité* and *Libération*. On the contrary, the right-wing tabloids like the *Sun* or the *Daily Mail* have no equivalents in France. However, the conservative *Le Figaro* could be compared to *The Telegraph* in the UK. Nevertheless, it was considered that the political systems were radically different. France is a republic and the UK a constitutional monarchy. Newspapers defining themselves as 'liberal,' 'conservatist,' or 'communist' are not cross-culturally equivalent. This is why political stances are not a criterion for comparing dataset. Moreover, conducting political comparisons in addition to cultural comparisons would complexify the analysis in an unnecessary way because the question of child sexual abuse is political, but cross-parties. What is at stake in this thesis is not to understand the political stance on incest, but how incest is defined in the press, as it may reflect how incest is perceived in the society

as a whole. Thus, the distribution of articles is calculated but only to provide a general overview of the data.

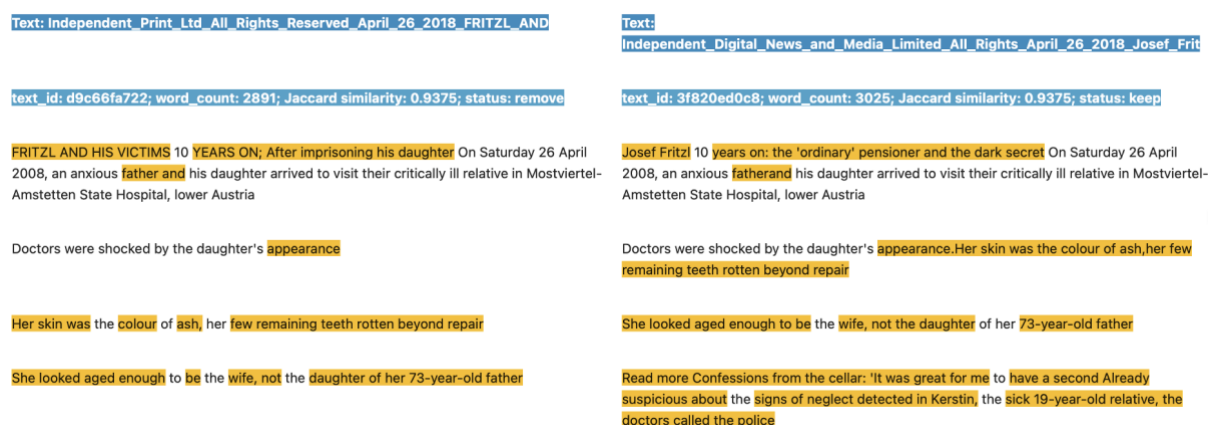
- (iii) *Finding a comparable search term.* Finally, a similar search term had to be found to build comparable corpora. The first search query was to collect all the articles with the lemma *incest** in both English and French, as it could include the English and French spellings *incest/inceste* and inflectional forms like the adjectives *incestuous/incestueux*. The search query *incest** retrieved 2,855 articles from the British newspapers, amounting to 3,192,777 words, and 3,541 articles from the French newspapers, amounting to 1,934,320 words. However, some articles were duplicates despite setting “Group Duplicates” to “moderate” on Lexis Nexis. Therefore, another tool was used to remove the duplicate articles in the corpora, as explained in the following section.

3.2.Sorting data

In order to discard the duplicate articles from the data, the Deduplication Tool developed on Jupyter Notebook by the University of Sydney (Asr et al., 2021; Broder, 2000) was used. The software allows the researcher to see which articles are duplicates and remove them. The software uses Python codes to calculate the percentage of similitudes between newspaper articles. Some articles are very similar because they are published twice, three times or even more times, undergoing editing after publication to correct wording or punctuation. Therefore, articles are identical except for the change of a comma, or almost identical as some words may have been changed or added. The aim is to avoid duplicates that would skew results in terms of frequencies.

Figure 4.3 shows how Jupyter notebook displays the two texts next to each other. This allows the researcher to examine both versions and choose whether to remove both, one, or keep both, depending on the research purposes. The example illustrates that such changes in wording are not innocent. As it was sometimes impossible to determine which version had been published first on the same day, it was decided to let the software do the deduplication itself in order to avoid tampering with the data unevenly. Since the aim of this study is to analyse the representation of incest, duplicates were removed.

Figure 4.3 Display of duplicate articles on Jupyter notebook



The software highlights in yellow the parts that differ. However, it shows the limit of the software as some parts are highlighted because they are not recognised by the software as identical because they appear on a different line. For instance, the line “Her skin was the colour of ash, her few remaining teeth rotten beyond repair” is on the next line in the left version, but closely linked to another sentence without space in the right version. However, this observation does not tamper with the results as the non-highlighted parts in white are detected more than the edited parts highlighted in yellow.

At the end of the process, 294 deduplicated articles were removed from the English corpus (corpus EN). After this process, Corpus EN comprises 2,561 articles, amounting to 2,882,435 words. In the French corpus (corpus FR), 192 deduplicated articles were removed through the same process. Corpus FR comprises 3,346 articles, amounting to 1,839,801 words. In total, it means that close to 10% of the words in the English corpus and close to 5% of the words in the French corpus were duplicated. This gives an idea of how skewed results may be when such duplicates are numerous in a corpus. This invites us to reflect once again on the challenges when a corpus is built. Despite the numerous computational tools available, it shows that the results always need to be considered with caution. For this reason, the analysis will be conducted using relative statistics on Sketch Engine, providing comparable frequencies.

3.3. Data distribution

This section presents the data distribution of the corpora. Corpus EN consists of 2,561 articles amounting to 2,882,435 words, while corpus FR comprises 3,346 articles amounting to 1,839,801 words. Figure 4.4 shows the number of articles in each corpus, and Figure 4.5 the number of words. As noticed, the corpora are uneven, with the English corpus counting approximately one million words more. Unexpectedly, there are more articles but fewer words in the French corpus than in the English corpus. It means that, on average, an English article comprises 1,125 words, while a French article contains 549 words.

Figure 4.4 Number of articles in corpus EN and corpus FR

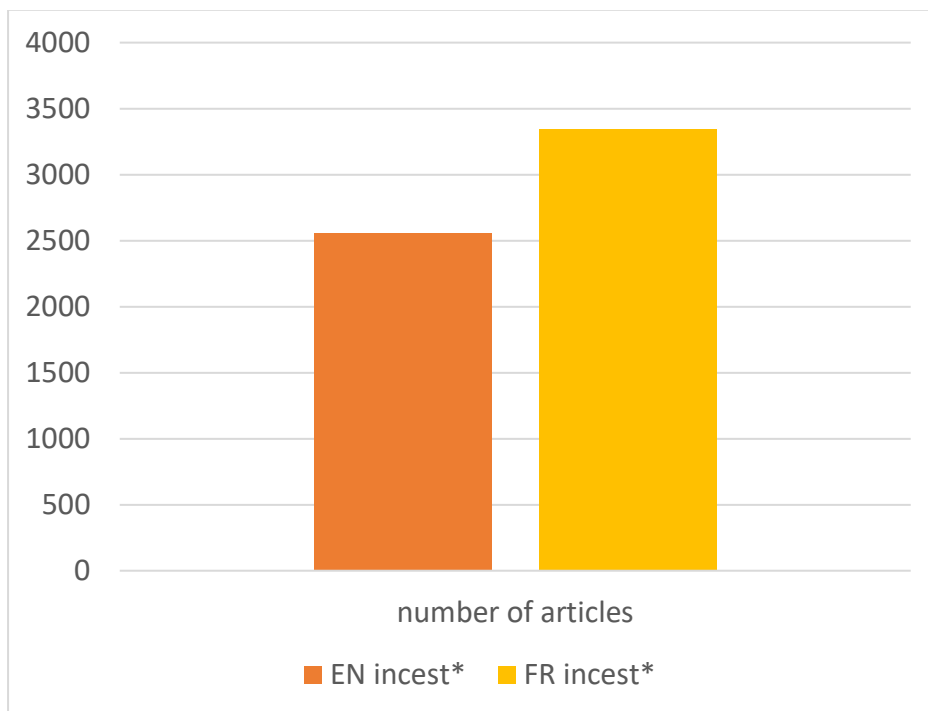
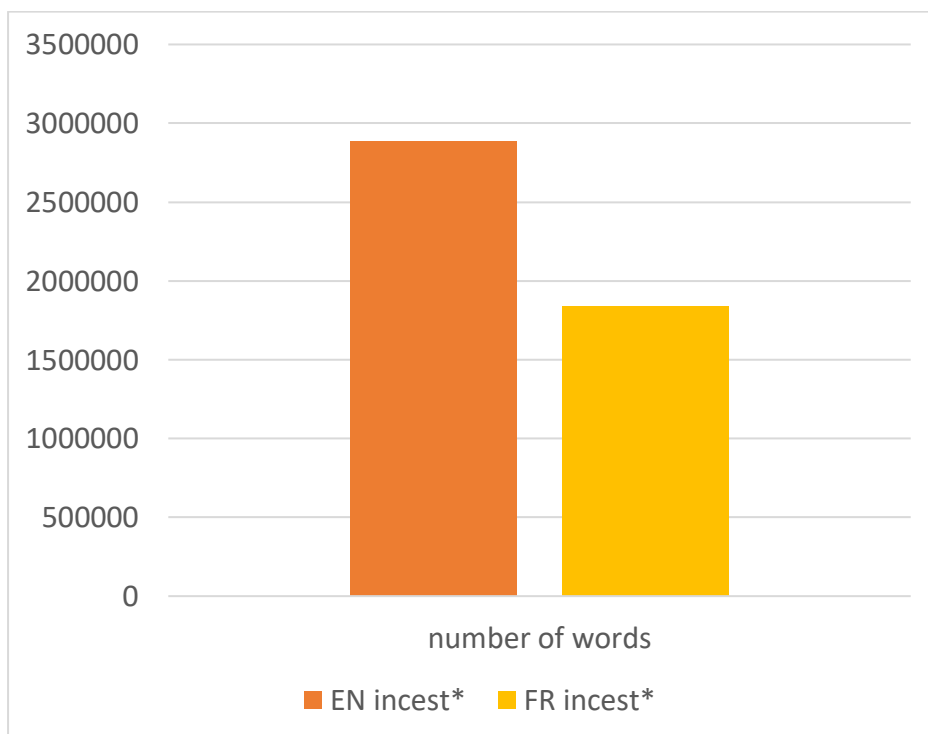


Figure 4.5 Number of words in corpus EN and corpus FR



Each corpus was then broken down into annual segments from 15 October 2017 to 14 October 2022. Since the data was collected at the end of the year 2017, articles published this year represent only 3% of the whole corpus in French and English, as seen in Figures 4. 6. and 4. 7. The percentages indicate the proportion of articles published over the six annual segments. Furthermore, the year 2022 is not entirely included as it ends on 14 October, but interestingly, the articles published in this year compose the largest proportion in the English corpus (28%)

as seen in Figure 9. The reason for this is explained in detail in the following chapter. In the French corpus, the year 2021 represents 40% of the corpus (Figure 4.6). This may be explained by the publication of the novel *La Familia grande* in January 2021, triggering a high-profile scandal (see Chapter 5).

Figure 4.6 Corpus FR breakdown per year

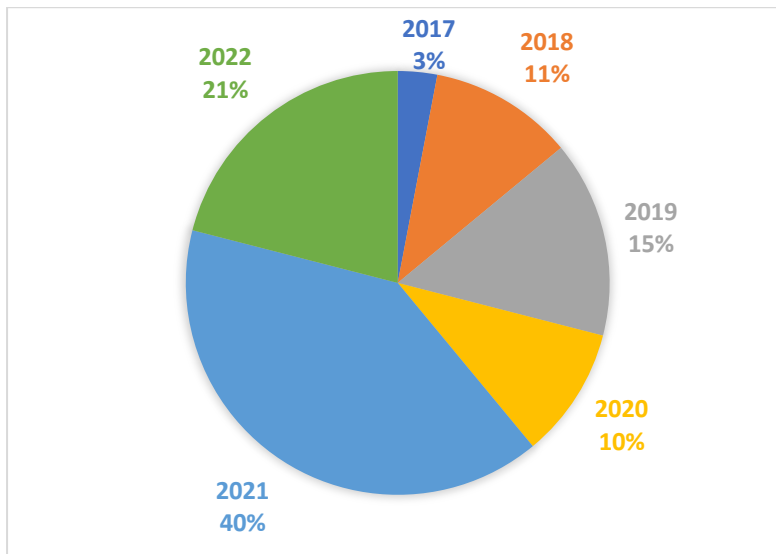
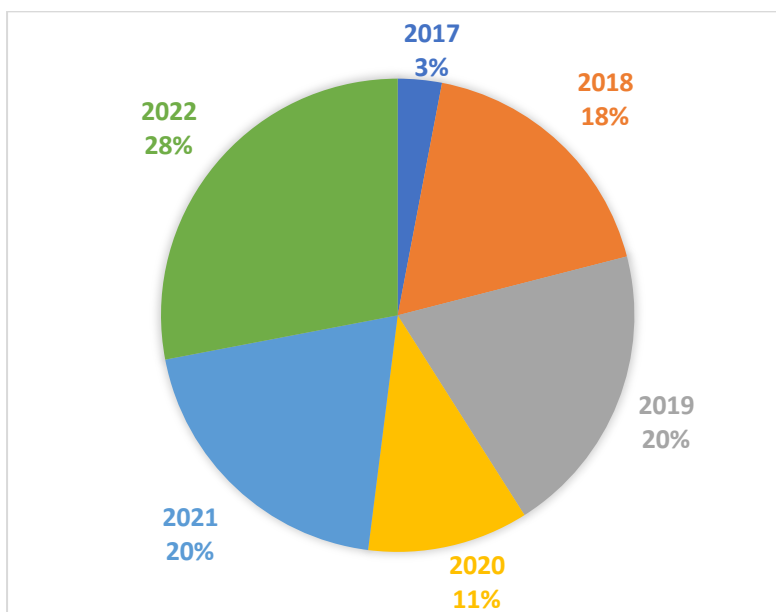


Figure 4.7 Corpus EN breakdown per year



Finally, for a fine-grained timeline, the corpus was broken down into monthly segments for both the French corpus (Figure 4.8) and the English corpus (Figure 4.9). Both figures indicate the number of articles published each month in raw numbers. As observed, different events involving the lemma *incest** were considered newsworthy at different time periods and for different reasons between the two countries. It should also be noted that Corpus FR reaches a peak of 235 articles in January 2021, likely related to the high-profile scandal that month. Such a peak is not reached in the English corpus, where the maximum number of articles

published with the lemma *incest** is 143 in May 2019. Overall, the publication rate is much lower in the British press than in the French press.

Figure 4.8 The monthly breakdown of articles in the French corpus

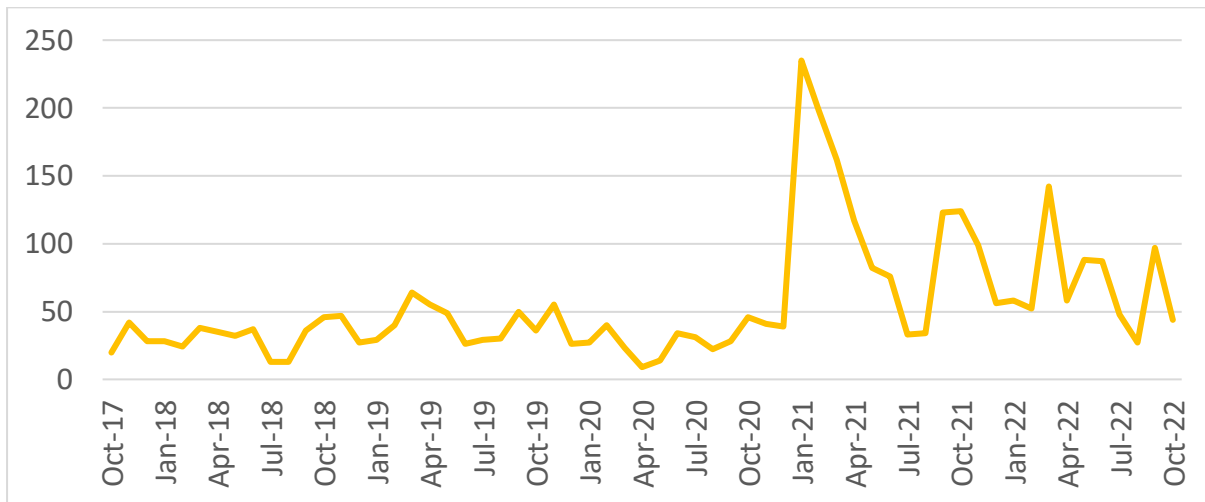
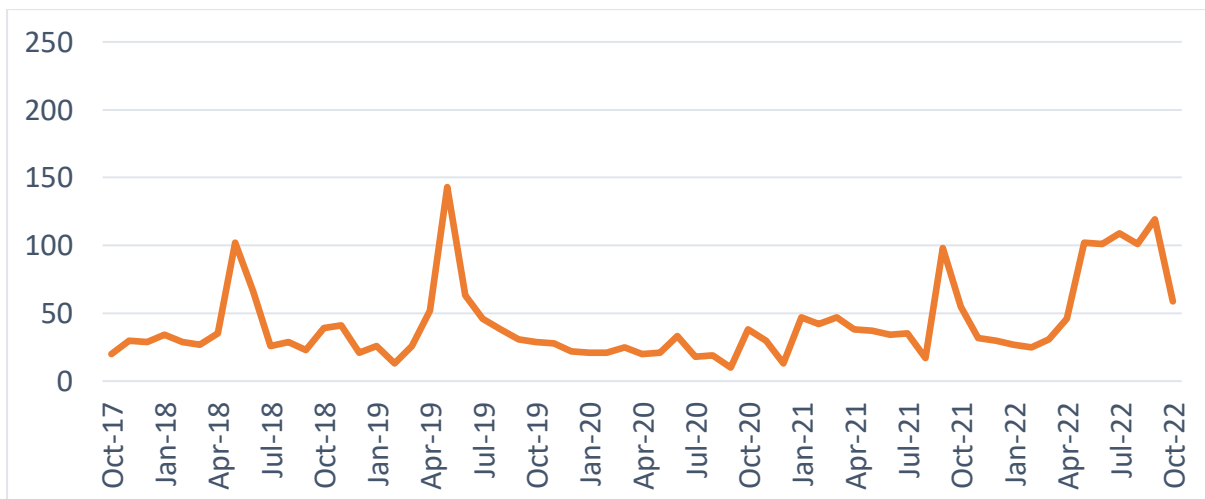


Figure 4.9 The monthly breakdown of articles in the English corpus



Finally, Figure 4.10 shows the proportions of articles between the regional and national newspapers in the French corpus, while Figure 4.11 shows the proportions of articles between the tabloids and broadsheets in the English corpus. Different colours are applied to the graphs to emphasise that the media landscape is not cross-culturally comparable, as discussed above. However, it is worth highlighting that the French regional newspapers published three times more articles involving *incest** than the national press. This confirms that the French regional press has a higher circulation rate than the national press. Furthermore, this suggests that incest is primarily framed in local news. In the English corpus, broadsheets published five times more articles than tabloids, indicating a sharper interest in incest in this type of press.

Figure 4.10 Corpus FR distribution

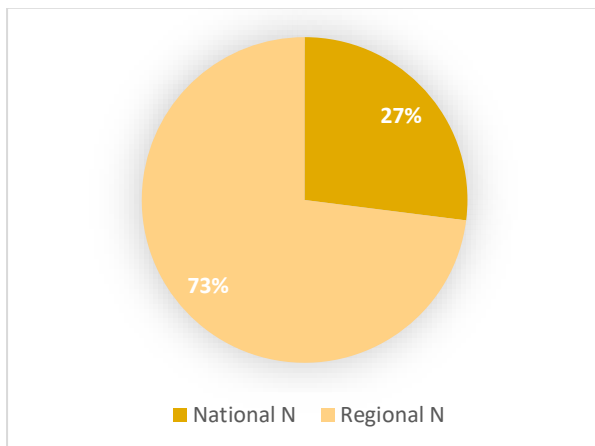
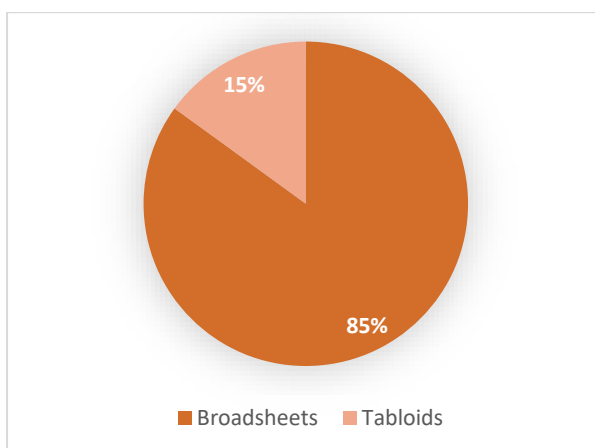


Figure 4.11 Corpus EN distribution



4. Summary

This chapter has defined the analytical framework of this research, addressing methodological challenges related to cross-linguistic CADS and discursive absence. The decisions made for building the corpora and addressing these challenges were explained to uphold transparency and replicability. This chapter argues that a novel methodological solution can be found in the double triangulation combining two methods and two linguistic datasets. The corpus was presented through graphs and diagrams that suggest cultural differences in news values for selecting events involving the lemma *incest**. The following Chapter 5 can now delve into the analysis of linguistic taboos surrounding incest in the French and British press between 2017 and 2022. The methodological framework is applied to contrast the representations of the incest taboo in the French and English languages.

Chapter 5: The Representation of the Incest Taboo in the French and British Press

Le tabou, dans notre culture, ce n'est pas le viol lui-même, qui est pratiqué partout, c'est d'en parler, de l'envisager, de l'analyser.

Triste Tigre (*Sinno, 2023, p. 24*)

What is taboo in our culture is not the rape itself, which is practiced everywhere, but speaking about it, considering it, analysing it.

After providing the theoretical and methodological frameworks of this thesis, this chapter applies cross-linguistic CADS to investigate the representation of the incest taboo in the French and British press from 15 October 2017 to 14 October 2022. In addition, this chapter aims to reflect on discursive absence. For this purpose, the French and English corpora, built with the same lemma *incest**, are contrasted to observe their similarities and differences. This chapter sheds light on the strengths and limitations of conducting a cross-linguistic analysis as “comparing news reporting in two languages adds an obvious complication, in that the corpora cannot be directly compared” (Hunston, 2006, p. 214). In addition, the definitions of the incest taboo, commented in Chapter 2, that perceive incest either as marital prohibition, or intrafamilial sexual abuse, or consented intrafamilial relationship, are discussed further in this chapter. Before delving into the analysis, it is worth noting that this chapter attempts to address the cross-linguistic challenges mentioned previously, specifically by making the French results accessible to non-French readers. To this end, the French results are translated into English in an added column or in parentheses in the tables. Regarding the diagrams downloaded from Sketch Engine, the data is translated in the body of text as the screenshots from the software could not be modified. Nevertheless, the linguistic patterns in the French corpus, and their similarities with and differences from the English corpus,⁵ are distinctive enough to be perceived even by a brief perusal, as the following sections demonstrate.

⁵ A distinction is constantly made between *English corpus* and *British press*. The *English corpus* refers to the corpus built with the search query, while the term *British press* is used to draw some general conclusions. The same distinction is made between the *French corpus*, referring to the data collected for this thesis, and the *French press*, mentioned when some conclusions can be drawn.

1. A cross-linguistic analysis

1.1. The topicality of each corpus

This section provides an overview of both corpora, aiming to understand how incest is represented in the French and British press (2017-2022). For this purpose, the frequency list of each corpus was generated on Sketch Engine. A frequency list indicates the most frequently occurring linguistic items in a corpus, such as lexical nouns, adjectives, adverbs, or prepositions. Thus, the most frequent lexical nouns can give insight into the topicality of the corpus. The topicality refers to aboutness of the corpus, in other words, what the corpus is about, and this is often determined through keywords (Baker, 2023; Scott & Tribble, 2006; Tranchese, 2023). However, as mentioned in the previous chapter, analysing keywords between two corpora in different languages raises cross-linguistic challenges (Taylor, 2014; Vessey, 2013). Normally, the corpus under investigation is contrasted with a reference corpus to determine its keywords. However, when dealing with two corpora in different languages, each corpus is contrasted with a different reference corpus, making cross-corpora comparison challenging. As each corpus is compared against a different reference corpus, the keyword lists necessarily differ. Furthermore, the keyword lists, despite being translated, cannot be directly compared from one language to another.

Therefore, another approach is used to analyse the topicality of each corpus in this thesis. The frequency lists of both corpora are examined to observe the most frequent lexical nouns in each of them. In doing so, the most frequent terms may indicate the aboutness of each corpus. As such, the frequency list indicates the most frequently occurring terms, suggesting how incest is framed in each press. Nevertheless, keyword lists are later generated to triangulate the findings. First, Table 5.1 displays the twenty most frequently occurring nouns in the French corpus. Cut-off points for such analysis are often arbitrary and cognitively biased in choosing round numbers (Baker, 2023; see Chapter 4). In this case, the number twenty was chosen due to the findings of the English corpus where the term *incest* is 20th in the list, as explained below.

Table 5.1 Top twenty lexical nouns in the French corpus

| Number | Noun | Translation | Raw frequency | Frequency per million |
|--------|-----------------|----------------------|---------------|-----------------------|
| 1 | an | year | 10,221 | 4,647.25 |
| 2 | enfant | child | 5,384 | 2,447.98 |
| 3 | <i>victime</i> | <i>victim</i> | 5,294 | 2,407.06 |
| 4 | inceste | incest | 3,907 | 1,776.42 |
| 5 | femme | woman | 3,657 | 1,662.75 |
| 6 | deux | two | 3,653 | 1,660.93 |
| 7 | fait | fact | 3,337 | 1,517.26 |
| 8 | <u>fil</u> | <u>girl/daughter</u> | 3,263 | 1,483.61 |
| 9 | <i>violence</i> | <i>violence</i> | 3,203 | 1,483.61 |
| 10 | viol | rape | 2,977 | 1,353.57 |

| | | | | |
|----|------------------|------------------|-------|----------|
| 11 | <u>père</u> | <u>father</u> | 2,942 | 1,337.66 |
| 12 | homme | man | 2,877 | 1,308.10 |
| 13 | h | hour (a.m./p.m.) | 2,609 | 1,186.25 |
| 14 | cour | court | 2,437 | 1,108.05 |
| 15 | <u>famille</u> | <u>family</u> | 2,349 | 1,068.04 |
| 16 | <u>mère</u> | <u>mother</u> | 2,337 | 1,062.58 |
| 17 | année | year | 2,262 | 1,028.4 |
| 18 | vie | life | 1,968 | 894.80 |
| 19 | jour | day | 1,917 | 871.61 |
| 20 | <i>agression</i> | <i>assault</i> | 1,873 | 851.61 |

In the frequency list reported in Table 5.1, two semantic fields can be identified. The underlined words in the list are related to family, with the occurrences of family members such as *père* (father), *mère* (mother), and *filles* –that has a double meaning in French as it may refer to a *girl* or a *daughter*, along with the term *famille* (family). The second semantic field, in italics in the table, relates to violence, with terms like *victime* (victim), *violence* (violence), *viol* (rape), and *agression* (assault). The word *incest* is the fourth most frequent term in the French corpus, indicating that the topicality is indeed about incest, framed in violence and a family environment. These preliminary observations suggest that incest is primarily represented as intrafamilial sexual abuse in the French press. Furthermore, the term *child*, the second most frequent term, highlighted as it is a frequent word in both corpora, suggests a particular focus on child sexual abuse. Nevertheless, this initial assumption will be further investigated to have a more precise understanding on how incest is represented in the French press.

The subsequent examination of the frequency list of the English corpus to determine its topicality offers unexpected results. The first most frequent term is *abortion*, and *incest* is only the twentieth most frequent term in the English corpus, as shown in Table 5.2.

Table 5.2 Top twenty lexical nouns in the English corpus

| Number | Noun | Raw frequency | Frequency per million |
|--------|--------------|---------------|-----------------------|
| 1 | abortion | 17,434 | 5,162.20 |
| 2 | woman | 10,896 | 3,226.30 |
| 3 | state | 7,769 | 2,300.40 |
| 4 | law | 7,225 | 2,139.32 |
| 5 | year | 6,374 | 1,887.34 |
| 6 | people | 5,800 | 1,717.38 |
| 7 | court | 5,789 | 1,714.12 |
| 8 | right | 5,492 | 1,626.18 |
| 9 | Trump | 4,690 | 1,388.71 |
| 10 | life | 4,429 | 1,311.42 |
| 11 | time | 4,429 | 1,301.65 |
| 12 | week | 4,376 | 1,295.73 |
| 13 | case | 3,856 | 1,141.76 |
| 14 | house | 3,607 | 1,068.03 |

| | | | |
|----|---------------|-------|----------|
| 15 | day | 3,572 | 1,057.67 |
| 16 | child | 3,468 | 1,026.87 |
| 17 | US | 3,424 | 1,013.84 |
| 18 | ban | 3,041 | 900.44 |
| 19 | rape | 2,972 | 880.01 |
| 20 | incest | 2,954 | 874.68 |

Thus, the findings from the frequency list indicate that the British press frames incest within the topic of abortion, and above all, within the topic of the abortion ban implemented in the US in 2022. A large semantic field (nine words out of twenty) refers to the American policy of banning abortion on a federal scale: *state, law, court, right, Trump, case, house, US, and ban*. As such, the British press focuses on the US internal policy where the bill “Roe versus Wade,” legalising abortion in the whole country in 1973, was overturned by the Supreme Court in June 2022 (Liptak, 2022). The ban gave the prerogative to the federal states to decide whether they allow abortion in their jurisdiction, thus limiting access to abortion. Even though Trump was no longer President of the US in 2022, his name appears in the 9th position in the list. He was held responsible for this backlash by appointing conservative judges at the Supreme Court. The terms *president* and *Biden* are respectively the 22nd and 23rd most frequent terms.

Therefore, these two frequency lists suggest that the coverage of incest differs in both countries. This contrastive analysis indicates that the semantic field of the family is comparatively absent in the English corpus, but the terms related to abortion are predominant, which is not the case in the French corpus. There are no terms referring to abortion or to the US in the French corpus, indicating a difference in framing the taboo topic of incest. The next step of the analysis involves analysing how abortion is related to incest in the British press. However, the findings of the frequency lists are first triangulated with those of the keyword lists. The French corpus is compared with the reference corpus FrTenTen23 and the English corpus with the reference corpus EnTenTen21 on Sketch Engine (Jakubíček et al., 2013). As commented in Chapter 4, these reference corpora are built through web scraping, providing an updated and varied language, even though the amount of data and the updates of both reference corpora differ. Notably, FrTenTen23 is two years more recent than EnTenTen21, but FrTenTen23 is half the size of EnTenTen21, amounting to 23.8 billion words and 52 billion words respectively. This demonstrates that the comparison of two corpora in different languages has inherent limitations. These limitations notwithstanding, the keyword list of the French corpus is reported in Table 5.3.

Table 5.3 Top twenty keywords of the French corpus compared to FrTenTen23

| Number | Keywords | Translation | Raw frequency | Frequency per million |
|--------|-------------------|-------------------|---------------|-----------------------|
| 1 | <u>inceste</u> | <u>incest</u> | 3,907 | 1,776.42 |
| 2 | <u>incestueux</u> | <u>incestuous</u> | 1,792 | 814.78 |
| 3 | Ciivise | Ciivise | 358 | 162.77 |
| 4 | Duhamel | Duhamel | 663 | 301.45 |
| 5 | réclusion | imprisonment | 559 | 254.16 |

| | | | | |
|----|------------------------|-------------------------------------|-------|----------|
| 6 | <u>viol</u> | <u>rape</u> | 2,977 | 1,353.57 |
| 7 | Kouchner | Kouchner | 454 | 206.42 |
| 8 | <u>attouchement</u> | <u>molesting</u> | 384 | 174.60 |
| 9 | prévenu | defendant | 891 | 405.12 |
| 10 | sociojudiciaire | sociojudicial* | 174 | 79.11 |
| 11 | <u>beau-père</u> | <u>stepfather/father-in-law</u> | 506 | 230.07 |
| 12 | IVG | termination of pregnancy | 427 | 194.15 |
| 13 | avortement | abortion | 1,108 | 503.78 |
| 14 | <u>pédocriminalité</u> | <u>pedocriminality*</u> | 169 | 76.84 |
| 15 | sursis | suspended sentence | 652 | 296.45 |
| 16 | assise | court | 1,153 | 524.24 |
| 17 | Roe | Roe | 178 | 80.93 |
| 18 | Angot | Angot | 178 | 80.93 |
| 19 | Mion | Mion | 138 | 62.75 |
| 20 | <u>belle-fille</u> | <u>stepdaughter/daughter-in-law</u> | 223 | 101.39 |

The keywords confirm the previous results from the frequency list: The French press discusses incest as a case of intrafamilial sexual abuse. The terms *incest*, *incestuous*, *rape*, *molesting*, *stepfather*, and *stepdaughter*, pertain to the same semantic field of incestuous abuse. Furthermore, the family names, *Duhamel*, *Kouchner* and *Angot*, all refer to incest. *Duhamel* is a former political scientist working in a prestigious Parisian university, who was accused by his stepdaughter *Camille Kouchner* to have sexually abused her twin brother in the book *La Familia grande* (Kouchner, 2021). *Angot* is the name of the writer Christine Angot who extensively writes about the incest she suffered from her father (Angot, 2013, 2017). Another linguistic pattern is the semantic field referring to justice, such as *réclusion* (imprisonment), *prévenu* (defendant), *sursis* (suspended sentence), *assise* (court), *sociojudiciaire* (sociojudicial) and *pédocriminalité* (paedocriminality). The last two terms are followed by an asterisk in the translation column in Table 5.3 because they are literally translated. They are French neologisms, not introduced in the English language. For instance, the neologism “sociojudicial” refers to a justice with social purposes. Furthermore, as touched upon in Chapter 4, *pédocriminalité* (“paedocriminality”) is a neologism created to substitute the term *paedophilia* that has a positive etymology with *philia* meaning *love* in Greek, whereas the term *criminality* leaves no ambiguity: sexually loving children is a crime. As the analysis will demonstrate, a translation equivalent is not only absent in English dictionaries, but also absent in the English corpus.

However, two terms related to abortion emerge: *avortement* (abortion), and the acronym *IVG*, standing for “Interruption Volontaire de Grossesse”, in other words, “termination of pregnancy,” meaning abortion. As such, the acronym *IVG* is an example of euphemising a taboo topic, like abortion, because it obscures the term (see Chapter 3). As such, *IVG* can be considered a euphemism of the orthophemism *avortement* (abortion) (Jamet, 2010). However, the terms related to abortion are less frequent than in the English corpus where most of the keywords (14 out of 20) directly refer to the abortion ban in the US (underlined in Table 5.4), thus confirming the previous findings reported above.

Table 5.4 Top twenty keywords of the English corpus compared to EnTenTen21

| Number | Keywords | Raw frequency | Frequency per million |
|--------|----------------------|---------------|-----------------------|
| 1 | incest | 2,954 | 874.68 |
| 2 | <u>abortion</u> | 17,434 | 5,162.20 |
| 3 | <u>anti-abortion</u> | 1,119 | 331.34 |
| 4 | <u>Roe</u> | 2,168 | 641.94 |
| 5 | incestuous | 486 | 143.90 |
| 6 | <u>foetal</u> | 448 | 132.65 |
| 7 | <u>near-total</u> | 257 | 76.10 |
| 8 | <u>pro-choice</u> | 381 | 112.81 |
| 9 | <u>overturn</u> | 1,230 | 364.20 |
| 10 | <u>foetus</u> | 265 | 78.47 |
| 11 | <u>Psaki</u> | 203 | 60.11 |
| 12 | rape | 3,401 | 1,007.03 |
| 13 | <u>Wade</u> | 1,499 | 443.85 |
| 14 | <u>Biden</u> | 2,726 | 807.17 |
| 15 | Rhaenyra | 152 | 45.01 |
| 16 | <u>DeSantis</u> | 246 | 72.84 |
| 17 | <u>parenthood</u> | 544 | 161.08 |
| 18 | <u>legalise</u> | 213 | 63.07 |
| 19 | Mastriano | 128 | 37.90 |
| 20 | Gainsbourg | 146 | 43.23 |

It is worth mentioning that the reference corpus EnTenTen21 was last updated in 2021, whereas the US Supreme Court overturned *Roe v. Wade* in June 2022. The keyword list displays what is typical in the corpus compared to the reference corpus. Therefore, since the reference corpus has been updated before 2022, there are necessarily more keywords related to the abortion ban. Conversely, the French reference corpus was updated in 2023, implying that there may be mentions of the abortion ban in the US in the collected data, mitigating the topicality of the abortion ban in the French corpus *incest**.

Nevertheless, the keyness analysis corroborates the frequency analysis. The keyword lists confirm that the primary focus of the French press on incest is to describe the topic as intrafamilial sexual abuse, whereas the focus of the British press on incest is embedded in the topic of abortion. Therefore, the incest taboo seems represented in different ways in the French and British press. These findings will be further investigated with a more fine-grained analysis carried out below. However, before considering the suggested differences, the following analysis focuses on the similarities between both corpora. It was noticed, and highlighted in Tables 5. 1. and 5. 2., that seven words are frequently occurring in both corpora: *woman, year, court, life, day, child* and *incest*. In order to determine the specificity of each corpus, the collocations of *incest, woman, and child*, are analysed in the following section.

1.2. The collocations of three similar lexical nouns

Corpus linguistics often studies differences rather than similarities in the data, due to cognitive bias (Baker, 2023; Taylor, 2013; Taylor & Marchi, 2018). Researchers tend to overreport what changes and differs rather than what remains the same (Baker, 2010). As such, corpus linguistics is inherently comparative (Partington et al., 2013; Partington & Marchi, 2015). However, it should not be interpreted as a sole interest for difference, but also for similarity. Indeed, corpus linguistics involves the observation of typical and recurring linguistic patterns (Hunston & Francis, 2000). Moreover, overlooking similarity to report difference is a methodological shortcoming, or “blind spot,” since “rather than aiming for a 360-degree perspective of our data, we are actually starting out with the goal of achieving only a 180-degree visualisation” (Taylor, 2013, p. 83).

Thus, the analysis in this subsection brings into focus the lexical nouns found both in the French and English frequency lists, that are *year: an/année*, *child: enfant*, *woman: femme*, *court: cour*, *life: vie*, *day: jour*, and *incest: inceste*. The aim is to determine whether these nouns indicate similarity between the press coverages, or whether the similarity is merely superficial. The preliminary results suggest that the coverage of the incest taboo differs between the French and British press. To this end, the collocations of three frequent lexical nouns in both corpora are examined: *incest*, *child* and *woman*. The other terms are discarded for frequency and semantic reasons. For instance, the term *year* can be translated by *an* and *année* in French, both of which frequently occur in the French corpus. Thus, it would skew the results to consider two different French words for one English word. Additionally, the terms *an* and *année* are used in different contexts, as *année* emphasises duration, contrary to *an*, which often indicates age. Ultimately, Sketch Engine only provides patterns with verbs for *année*, and not for *an*. Therefore, the grammatical patterns are not only challenging to compare between the French and English words but also between the two French words.

Therefore, the collocation analysis focuses on *incest: inceste*, *child: enfant*; and *woman: femme*. As mentioned earlier, collocations refer to frequently co-occurring words that are “especially interesting to investigate, as they can point to the salient ideas associated with a particular phenomenon” (Jaworska & Krishnamurthy, 2012, p. 6). As such, collocations are computed through statistical measures. The association measures yield different results depending on the criteria selected for retrieving collocations (Brezina, 2018). The Mutual Information (MI) measure retrieves the collocates that are exclusive, even though they rarely occur with one another (Brezina, 2018; Tranchese, 2023). LogDice is another measure to calculate “the typicality of a collocation (as opposed to its frequency), so it tends to favour reasonably low frequency collocates although usually not as low as the ones obtained using MI” (Baker, 2023, p. 142). The metrics LogDice is available on the Word Sketch tool developed by Sketch Engine, on which this research relies. Word Sketch is a useful tool to visualise “the word’s grammatical and collocational behaviour” (Sketch Engine). In other words, “it not only provides a list of words that frequently co-occur with the node word, but it also provides information about the grammatical relation that these collocates have with the node word”

(Tranchese, 2023, p. 153). Furthermore, Word Sketch provides a colourful visualisation that enhances the most frequent collocational and grammatical patterns through the size and positions of the words. As Word Sketch relies on the measure LogDice, it represents the most typical patterns in a bigger font and close to the node term placed at the centre of the circle. The less typical patterns are visually represented at the further end of the circle and in smaller font. Thus, collocates like the adjective *other* or the verb *be* are frequent but not typical, because they often collocate with many terms. As such, they are visualised in a bigger font to represent the frequency but are positioned at the further end of the circle, as they are not typical. These explanations will become clearer with the following diagrams, Figures 1-5. For the sake of diversity, different types of collocational and grammatical patterns are presented for the three lexical nouns.

With this in mind, the collocates of *incest*, *woman*, and *child* in both languages are examined to investigate how these terms are discursively constructed in each press coverage. Before presenting the characteristic patterns of *woman* and *child*, the focus is on *incest*. The lexical and grammatical collocates of *incest/e* are observed in both corpora with Word Sketch. It is worth noting that *incest* occurs 2,952 times in the English corpus, and *inceste* occurs 3,907 times in the French corpus. In relative numbers, it represents 0,10% of the English corpus, and 0,21% of the French corpus, meaning that the term is twice as frequent in the French corpus as it is in the English corpus. The following figure, Figure 5.1, contrasts the findings for the term *incest* between the two corpora. In this case, the collocates are lexical nouns that follow *incest* after the conjunctions *and/or*, including punctuation like the comma. Thirteen collocates are displayed on the diagrams for clearer visualisation. The full lists of collocates are presented and analysed in the following section. From a preliminary observation, it can be noticed that the most frequently occurring collocate of *incest* is *rape: viol*, in both languages. It indicates that *incest* is associated with sexual abuse. A closer observation reveals that *incest* collocates only with terms having negative meanings in both languages, with the exceptions of the collocates *life* and *health* in the English corpus. Except these two words, *incest/e* is associated with terms pertaining to extreme violence: *rape: viol*, *violence: violence*, *abus: abuse*, *pédophilie: pedophilia*, *pédocriminalité* (pedocriminality), *harcèlement* (harassment), *danger*, *agression* (assault), *délit* (offence), *parricide* (parricide), *murder*, *trafficking*, *bestiality*, *threat*, *assault*.

However, three points of divergence should be highlighted: (i) there are closer interconnections between the node term *incest* and its collocations in the French corpus, (ii) *inceste* often occurs with the terms *viol* (rape) and *violence* in the French corpus, while in the English corpus, *incest* is mainly related to *rape* alone, (iii) there are two terms referring to paedophilia in French: *pédocriminalité* and *pédophilie*, the first term being a neologism to emphasise the crime committed by abusing children. Added together, they are the third most recurrent terms collocating with *incest*, (iv) *abnormality* (115) is the second most recurring term in the English corpus, occurring in relative numbers five times more frequently than its counterpart *malformation* (15) in the French corpus.

Figure 5.1 *incest* and/or...* in the French and English corpora

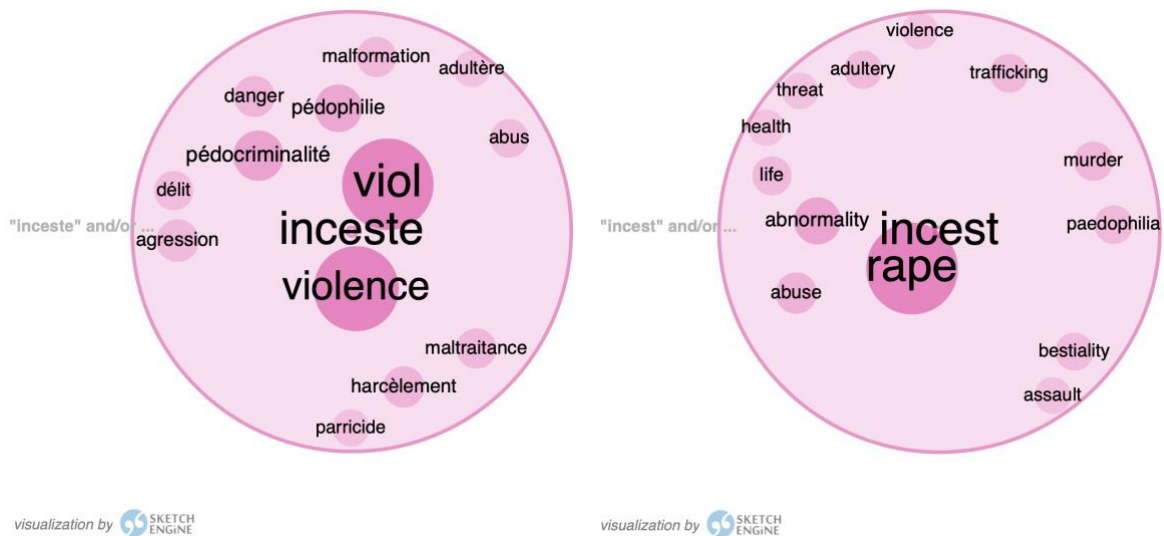
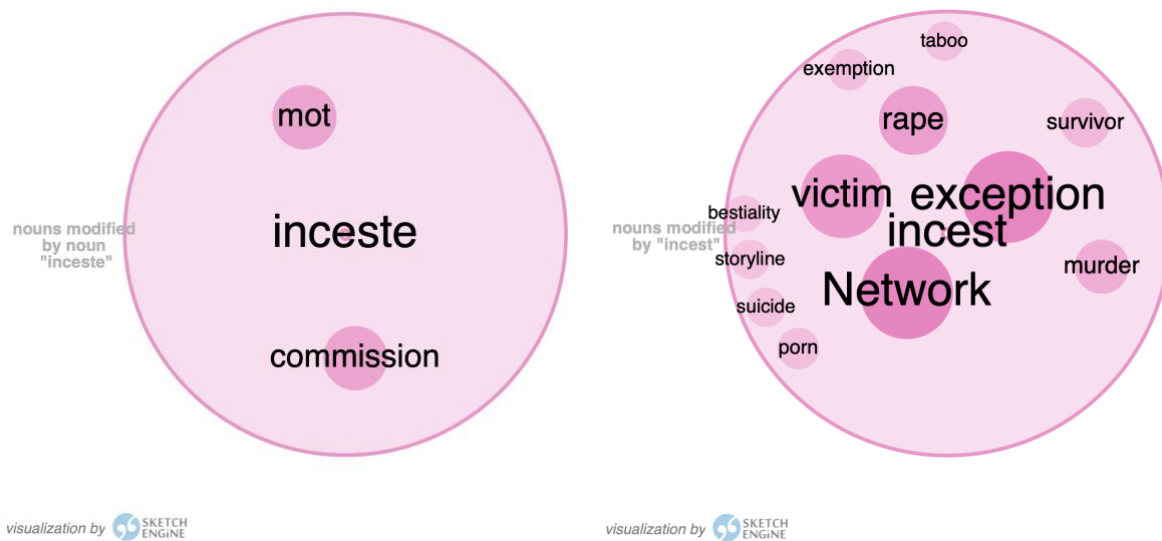


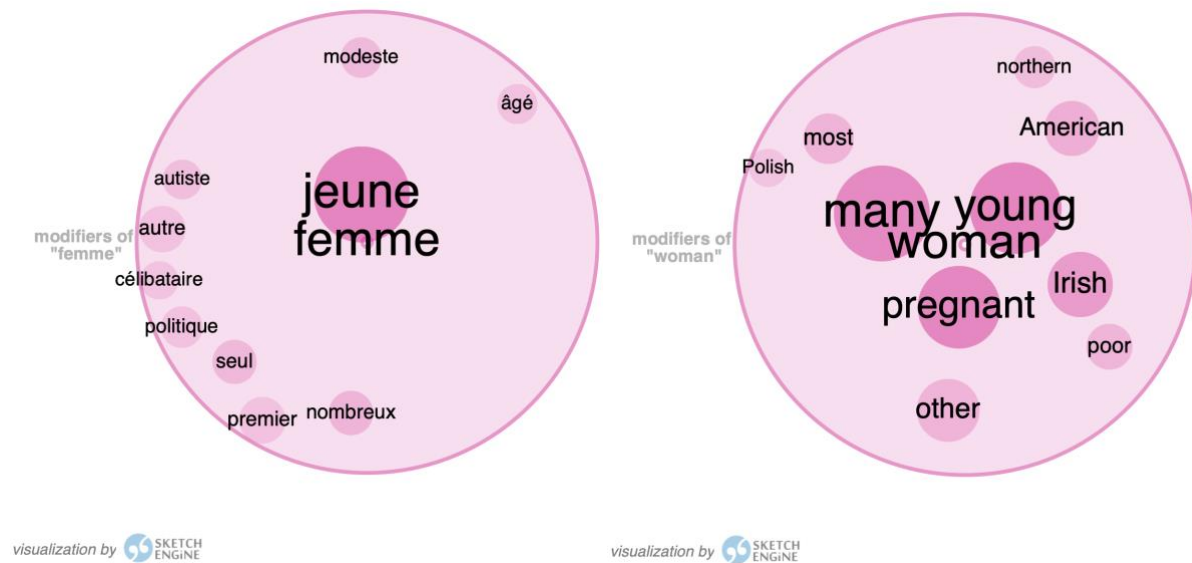
Figure 5.2 illustrates the challenges encountered in cross-linguistic analysis. It displays the nouns modified by *incest/e* and highlights the complexity of contrasting two languages inherently different. In the French language, a noun is modified by another noun with prepositions, unlike in English, where nouns can directly modify another noun, without using a preposition. Therefore, Sketch Engine does not easily analyse nouns modified by *incest/e* in the French corpus. As can be observed, the French diagram is almost empty, whereas the English diagram displays several terms, many of which pertain to acts of violence: *victim*, *rape*, *murder*, *survivor*, *suicide*, *bestiality*. However, the English diagram has limitations as well. Some terms are not directly modified by the noun *incest*, as indicated by the software, but are closely preceding *incest* with a comma, like for *murder* and *rape*, always associated with incest through a comma like in “*incest, murder*” and “*incest, rape*”. The software overlooks that these terms collocate in an enumeration. Conversely, other terms like *victim* or *survivor* are directly modified by the term *incest*, as in “*incest victim*” and “*incest survivor*”. Thus, the software does not make this distinction in the grammatical patterns, revealing a limitation.

Figure 5.2 Nouns modified by *incest/e* in the French and English corpora



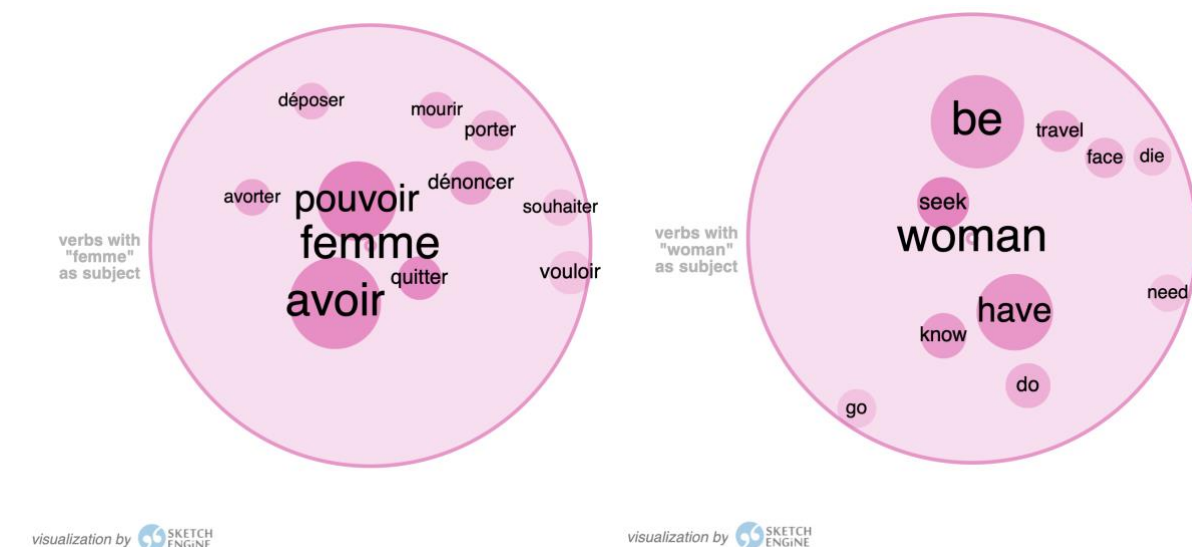
After briefly examining the collocates of *incest/e*, the analysis focuses on the collocational patterns of *woman: femme* and *child: enfant*. These are the two most frequent terms appearing in both corpora, suggesting that they are closely related to the topic of incest in the French and British press. In frequency per million, the term *woman* occurs 3,226.30 times in the English corpus, while *femme* appears 1,662.75 times in the French corpus. Thus, the term *woman* is twice as frequent in the English corpus as in the French corpus, implying that the focus on women is sharper in the British press. Furthermore, Figure 5.3 shows how *woman* has more typical collocates in the English corpus. However, a closer examination indicates that these collocates relate to the topic of abortion: *pregnant* (181), with the names of countries where the right for abortion is limited, such as *Irish* (111), *American* (78), *Polish* (36), and *Northern* (44), often used as in the phrase “Northern Irish women.” As observed earlier, such focus on abortion is specific to the English corpus. The modifiers of *femme* in the French corpus are less typical and more diverse as they refer to different topics without distinctive semantic fields: *nombreux* (numerous), *seul* (alone), *modeste* (modest), *autiste* (autistic), *politique* (political), *autre* (other), *premier* (first), and *célibataire* (single). As noticed in other studies (Gerding Speno & Aubrey, 2018; Jackson & Tinkler, 2007), the term *woman: femme* is closely associated with the notion of *youth*. In both corpora, the adjectives *young* (225) and *jeune* (416) are the most frequent collocates.

Figure 5.3 Modifiers of *woman: femme* in the French and English corpora



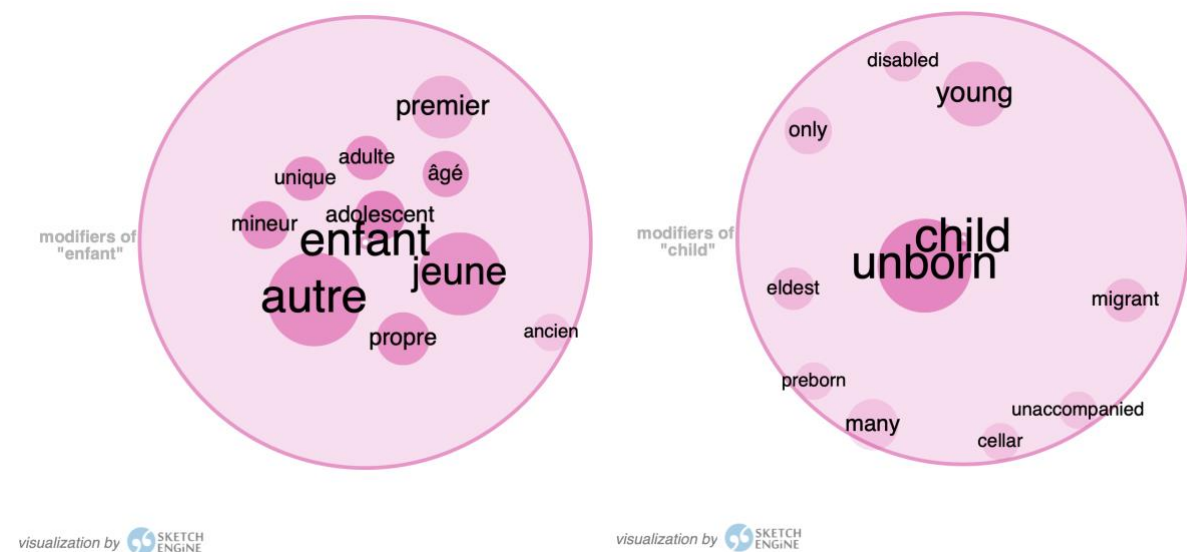
An additional interesting linguistic pattern to observe is the most frequently co-occurring verbs with *woman: femme* as subject, as shown in Figure 5.4. Notably, the term *woman* seems more a passive subject in the English corpus than in the French one as the most frequent verbs imply powerlessness: *seek* (169), *face* (62), *die* (48), and *need* (43). Conversely, the collocating verbs in the French corpus indicate a more active portrayal of women: *pouvoir* (can) (23), *quitter* (leave) (8), *dénoncer* (denounce) (8), *avorter* (abort) (6), *déposer* (prosecute) (6), *porter* (carry/prosecute) (7), *souhaiter* (wish) (6), and *vouloir* (want) (8). All the verbs collocating with *woman* as their grammatical subject in the English corpus refer to abortion, including the verb *travel* (71), seemingly portraying woman in a more active way, but used when women are constrained to travel to another state or country to access abortion. Thus, despite being the top five most frequent terms in both corpora, the terms *woman: femme* are used in different contexts, suggesting superficial similarities that the collocation analysis uncovers.

Figure 5.4 Verbs with *woman: femme* as subject in the French and English corpora



The final frequent term under investigation is *child: enfant*. In frequency per million, this term occurs 1,026.87 times in the English corpus and 2,447.98 times in the French corpus. Thus, the term *enfant* is more frequent than *child* between the two corpora – unlike the terms *woman: femme*, occurring more frequently in the English corpus than in the French corpus. This suggests that the French press places emphasis on children, whereas the British press focuses on women, indicating that each press covers incest from different angle based on their selected news values. This difference in angle is further illustrated by examining the collocates modifying *child: enfant* in both corpora, as shown in Figure 5.5. Similar patterns to those observed for *woman: femme* can be observed in the co-occurring modifiers of *child: enfant*. Figure 5.5 demonstrates that the British press primarily focuses on abortion with the main collocate being *unborn* (138). In contrast, the French corpus shows varied collocates, many of which refer to age, such as *adolescent* (teenager) (11), *mineur* (underage) (10), *adulte* (adult) (8), *jeune* (young) (39), and *âgé* (aged) (9). This seems to confirm that the British press focuses on children before they are given birth (in relation to abortion), while the French press reports different stages of children’s life. Thus, the British press frames the topic of incest within the context of abortion, a framing that does not seem to be present in the French press. The reasons why abortion is such a predominant theme in the British press are explored in the following subsection.

Figure 5.5 Modifiers of *child: enfant* in the French and English corpora



1.3. The abortion ban in the US

This subsection investigates how incest and abortion are linked in the English corpus. The frequency and keyword lists, along with the collocation analysis, reveal the same pattern: the British press primarily discusses the incest taboo within the context of abortion, and more specifically, with reference to the abortion ban in the US implemented in 2022. This decision let each state legislate the access to abortion. In January 2024, abortion was banned in twenty-one states (The New York Times, 2024). Unexpectedly, a large majority of articles in

the English corpus covered this news while using the term *incest** at least once. This is why the largest proportion of articles in the English corpus were published in 2022 (see Chapter 4).

In addition, the term *rape* co-occurs 1,739 times with the term *incest*, while the latter occurs 2,954 times in the whole corpus. This implies that more than half of the instances of *incest* are closely associated with *rape*. In order to observe how the two terms collocate, a concordance analysis is conducted on the grammatical pattern *incest and/or rape*. As outlined in Chapter 4, a concordance analysis consists in observing the noun in context. As such, the node term is displayed at the centre of the screen, highlighted in red on Sketch Engine, and preceded and followed by the sentence in which it appears. The concordances analysis facilitates the examination of the context in which the term is used. To this end, concordance analysis “is based on the close examination of text from a corpus,” relying on the researcher’s close reading, which can be a fastidious work. This is why different methods have been developed to conduct systematic and replicable concordance analyses (Baker, 2023, pp. 112–115). For instance, Baker (2023) mentions the concept of closure, developed by Sinclair (1999), that consists in analysing every thirty lines until no new patterns are found. However, every researcher should apply the method which yields the best results for them. For instance, Baker (2023) prefers to examine 100 lines “to identify the most common features around or uses of the search term as well as the less frequent features” (p. 113), but Tranchese (2023) prefers to select “a percentage, rather than a given number,” namely 10% of the concordance lines of her corpora to identify patterns (p. 155).

In this study, given the manageable size of the corpora, all concordances lines were ordered alphabetically on the left side to facilitate the visualisation of linguistic patterns. Quickly enough, a distinctive linguistic pattern emerged: the phrases “*with no exceptions for rape or incest*” or “*in cases of rape and incest*” were overwhelmingly repeated, as if newspaper articles were quoting the same sentence from an identical source. In order to quantify the frequency of these patterns, the N-grams are examined. The N-gram tool indicates clusters of words that are frequently occurring in the corpus (Hunston, 2006). The N-gram length is set at 5, meaning that all the occurrences of the identical clusters of five words are retrieved. The list of 5-grams confirms the previous observation: the phrases “*even in cases of rape or incest*” or “*no exception for rape or incest*” are overwhelmingly repeated in the corpus, as highlighted in Table 5.5.

Table 5.5 The fifteen most frequent 5-grams in the English corpus

| Numbers | 5-grams | Frequency |
|---------|--------------------------------------|-----------|
| 1 | cases of rape or incest | 167 |
| 2 | the life of the mother | 158 |
| 3 | exceptions for rape or incest | 155 |
| 4 | in cases of rape or | 153 |
| 5 | cases of rape and incest | 135 |
| 6 | in cases of rape and | 119 |
| 7 | no exceptions for rape or | 101 |
| 8 | to overturn Roe v Wade | 93 |

| | | |
|----|---------------------------------------|----|
| 9 | in the wake of the | 89 |
| 10 | the right to an abortion | 84 |
| 11 | even in cases of rape | 82 |
| 12 | to save the life of | 75 |
| 13 | result of rape or incest | 74 |
| 14 | exceptions for rape and incest | 74 |
| 15 | except in cases of rape | 66 |

Thus, Table 5.5 gives insights into the frequent co-occurrences of *incest* and *rape*. The British press reports the decision of the Supreme Court to ban abortion without exceptions for rape or incest. Hypothetically, if such an event in the US had not taken place, fewer articles would have been retrieved through the lemma *incest** in the English corpus. It suggests that the British press does not often employ the term *incest*. Nevertheless, the other instances of *incest* are worth exploring in the English corpus. Indeed, despite the predominance of the topic of abortion, incest is covered in other contexts in the British press. The following section conducts a more fine-grained analysis to examine the representation of the incest taboo in the French and British press.

2. The discourse prosody of incest in the French and British press

This section carries out a more fine-grained analysis of the representation of incest in the French and British press. In doing so, the study focuses on collocation and concordance analyses. The various terms co-occurring with *incest** have been analysed above, however, their concordance analyses give insight into the semantic preference and discourse prosody on the incest taboo. Semantic preference refers to the terms that frequently co-occur together. It means that some words prefer the company of other words, having either a more positive or negative meaning (Hoey, 2005; Morley & Partington, 2009; Partington, 2004). Indeed, words are always in relation with other words in “predictable combinations” (Stubbs, 2001, p. 3). However, as outlined in Chapter 4, Section 2, analysing semantic prosody across languages is challenging as it is influenced by language and culture, which implies that it is not easily transferable from one language to another (Grabowski, 2022; Vessey, 2013). For this reason, the following analysis takes steps to ensure replicable linguistic contrasts, even though it is acknowledged that such cross-linguistic comparisons are limited due to inherent differences in lexicogrammar structures. Therefore, this section examines the discourse prosody of each press regarding the incest taboo.

2.1. The eroticisation of incest

“This is incest, of course, but it’s also Greek myth.” The quotation is taken from an article published in *The Times* in December 2019 that covered the temporary exhibition on sin staged at the National Gallery in London. In this exhibition, different types of sins were artistically depicted. One painting caught the journalist’s attention: Bronzino’s *Allegory with Venus and Cupid* where Cupid and his mother Venus are kissing. The journalist gives the following description:

Bronzino's Cupid is adolescent and is embracing his mother, kissing her on the lips. His fingers are caressing a nipple that seems (shall we say?) interested. This is incest, of course, but it's also Greek myth and given the various ways and guises in which Zeus rapes his victims, a bit of consensual Cupid-on-Venus action seems less shocking.

This comment is representative of the eroticised depiction of incest in societies at large (Brey & Drouar, 2022; Harkins, 2009), and in the press more specifically. The journalist's description emphasises Cupid's action, namely that the child is the active agent of the kiss: "is embracing", "is kissing", "his fingers are caressing", "*consensual* Cupid-on-Venus action," which corresponds to the general perception of the Oedipal complex, implying that the children are the *subjects* of the desire, and not the *objects* of the desire (see Chapter 2). Needless to say, it would be a mistake to condemn pictural representations of incest painted centuries ago, but it is interesting to notice that such fantasised depictions are still an appealing focus in the press.

Therefore, this subsection aims to understand how incest is portrayed in the press. As observed above, the collocation analysis of the term *incest* reveals that its collocates almost exclusively relate to extreme violence. To some extent, and to further refer to the exhibition at the National Gallery, the collocates can be defined as other types of sins, as seen in Table 5.6.

Table 5.6 The collocates of *incest* in the pattern *incest and/or...* in the English corpus

rape (1,739), abnormality (115), **abuse** (47), **murder** (30), **trafficking** (26), **paedophilia** (25), life (28), **adultery** (22), **bestiality** (21), **threat** (13), **assault** (13), **violence** (13), health (13), **treason** (9), **danger** (9), **death** (9), **cannibalism** (8), **risk** (8), **sodomy** (7), **suicide** (7), emergency (7), pregnancy (7), **cruelty** (6), lawmaker (6), malformation (5), dragon (5), exception (5), abortion (6), sex (5), woman (7), thalidomide (4), **molestation** (4), infidelity (4), dancing (4), **racism** (4), child (4)

The fact that incest is an act of violence is clear for all victims, and any other victims of sexual abuse (Sinno, 2023). However, we can wonder in which context terms like *bestiality* and *cannibalism* co-occur with *incest*. The hypothesis is that incest may be imbued with monstrosity, as it is often reported for other types of sexual abuse like rape. Tranchese (2023) demonstrates that the term *rape* is mainly used in reference to "wartime rape" since women's responsibility is not called into question in the context of wars. However, whenever a celebrity is involved, like a politician or an athlete, the term *rape* is often substituted by the euphemism *alleged incident* (Ibid., p. 194-199). Therefore, *rape* is an orthophemism when violence is clearly stated but is euphemised when violence is alleged. It is easier to depict soldiers of the opposite camp as monsters than celebrities. However, monsters do not exist, even though their depictions are convenient for not blaming common humans (Lamy, 2023). Thus, it is of interest to take a closer look at the link between *incest* and its collocates connoting extreme violence. The aim is to investigate in which contexts they co-occur.

Unexpectedly, the concordance analysis reveals that incest is used in reference to fiction in almost all the occurrences. Two patterns can be pointed out: (i) incest is part of an enumeration, where it is mentioned among other violent crimes, sometimes ironic and downplaying incest (like *cat murder*), sometimes increasing the emotional charge (like *torture*), (ii) incest is portrayed in fiction, including pornography, series, and films. As such, incest is presented as a subplot of fiction, entrenched in imagination, and contributing to sordid stories. It represents incest in an imaginary realm, as seen in Table 5.7, where the terms referring to fiction and other crimes are highlighted.

Table 5.7 Sample of concordance lines of *incest* with the collocates *murder*, *cannibalism*, and *suicide* in the pattern *incest and/or...*

| N | Np | Date | co-text | node | co-text |
|---|-----|------------|---|----------------|--|
| 1 | Ind | 24/02/2019 | This is a director who, just under a decade ago, released his breakout film <i>Dogtooth</i> - a grim tale of isolation, | incest, | cat murder and DIY dentistry. |
| 2 | Tim | 10/05/2019 | <i>Game of Thrones</i> has battles, dynastic feuding, | incest, | murder , mayhem and extraordinary castles and palaces . |
| 3 | Ind | 27/04/2020 | Here's a sample of some of the videos and story categories available on the four most popular adult websites in 2020: mind control, slavery, violence; bestiality, nonconsensual sex , body modification, cruelty, drug, rape , reluctance, snuff, torture , young, blackmail , humiliation, | incest, | murder . |
| 4 | Tim | 21/09/2018 | The script then delivers a sinister biographical subplot involving | incest | and suicide . |
| 5 | Gu | 18/09/2019 | It was not a roman à clef but , given her Bloomsbury background, some readers inevitably sought biographical clues in its knotty storyline , with its intimations of | incest | and suicide . |
| 6 | Gu | 04/08/2020 | Murder , | incest, | cannibalism , a "singing anus" - and, of course, Divine scarfing down a freshly produced dog turd - my God, this film had it all . |

Therefore, incest is described as a fictional element plot, with a rich semantic field pertaining to contemporary fiction such as *director*, *film*, *writer*, *reader*, *story*, *storyline*, *script*, *tale*, *fairytale*, *novel*, *roman*, *satire*, *video*, *adult website*, *platform*, and *site*. Furthermore, incest is not frowned upon but rather enjoyed, as evidenced by positive comments in examples (1) and (3) like “**breakout film**” and “**most popular adult websites**,” or the journalist’s expression of wonder in example (6), “my God, this film had it all”. This fictional and consensual depiction

of incest is in line with Brey’s analysis (2022) stating that fiction displays eroticised representations of incest, featuring consenting adults in a transgressive but desired sexual relationship, instead of representing incest in its more prevalent form, namely as child sexual abuse (see Chapter 2). As such, eroticised representations of incest are widespread in contemporary fiction and exist along a continuum of pictorial depictions from previous centuries.

Moreover, the fact of mentioning incest in an enumeration of sordid acts contributes to staging incest within the fictional realm of extreme violence. Thus, incest is further marginalised from social reality by being portrayed in a fantasy universe. In addressing the question whether incest is framed in monstrosity, it can be argued that to some extent, it is. Not monstrosity in the sense that Tranchese (2023) noted, where rape is only acknowledged in instances of extreme violence (like in war), but rather monstrosity as depicted in fiction, which displays actual monsters. The terms *cruelty*, *torture*, *slavery*, and others, co-occurring with terms related to fiction, portray monstrosity in pure imagination. The figure of the monster serves as a convenient scapegoat to avoid confronting the prevalence of incestuous abuse. Indeed, the high prevalence of incest (as discussed in Chapter 2) reveals that sexual abuse within families extends beyond the realm of castles and palaces surrounded by dragons.

As a side note, other terms pertaining to extreme violence are used in the context of abortion. As mentioned above, abortion is a recurrent theme in the English corpus, complexifying the findings on incest. In Table 5.8, a sample of concordance lines demonstrate how *incest* co-occurs with *abortion*, along with the collocates *threat*, *trafficking* and *danger*.

Table 5.8 Sample of concordance lines of *incest* with the collocates *threat*, *trafficking*, and *danger* in the pattern *incest and/or...*

| N | Np | Date | co-text | node | co-text |
|---|-----|------------|--|---------------|--|
| 1 | Tim | 23/10/2020 | The ruling will permit abortion only in cases of rape , | incest | or a threat to the mother's health and life. |
| 2 | Ind | 07/08/2022 | It does not allow exemptions in cases where pregnancies were caused by rape , | incest | or human trafficking . |
| 3 | Gu | 29/05/2018 | The provision, enshrined in the constitution since 1983, gave "the unborn" equal rights to pregnant women , making abortion illegal even in cases of rape , | incest | and severe danger to the mother . |

It shows once more that incest is mainly mentioned within the topic of abortion in the British press. Furthermore, abortion is framed in a context of violence, where the conditions for abortion are exemplified through extreme cases where danger prevails. In example (1), the case takes place in Poland, where abortion is allowed when the mother’s life is *threatened*. Example (2) refers to the non-exception clause in the US for having access to abortion even in case of *human trafficking*. Finally, example (3) refers to Ireland’s abortion ban even if the

mother's life is in *danger*. In this table, the three lines represent three different countries, showing similar restrictions on access to abortion across them.

The following step is to contrast the findings. Hence, the collocates of *inceste* are analysed in the French corpus to determine whether similar fictional representations of incest exist in the French press as well. The same collocational pattern in the French corpus, i.e., *inceste and/or...*, is examined for this purpose. As observed in the English corpus, incest frequently co-occurs with terms denoting extreme violence, as highlighted in Table 5.9. It is worth noting that the term *sodomy* has not been highlighted but underlined (as in Table 5.6 for the English corpus). This decision comes from the fact that sodomy is merely a sexual act, even though it is disapproved and marked as immoral as it is mainly performed between two men, thus directly or indirectly referring to homosexuality, still widely rejected, mainly by religion (Jordan, 2000). For this reason, it cannot be highlighted on the same grounds as terms like *murder*, *cannibalism*, or *infanticide*. Regardless of considering acts like sodomy (or adultery) as moral or immoral, their co-occurrences with *incest* in both corpora indicate that their negative connotation contributes to the negative semantic prosody surrounding incest.

Table 5.9 The collocates of *inceste* in the pattern *inceste and/or...* in the French corpus

viol rape (351), **violence violence** (291), **pédocriminalité paedocriminality** (46), **pédophilie paedophilia** (36), **harcèlement harassment** (18), **danger danger** (17), malformation *abnormality* (15), **agression assault** (21), **délit crime** (13), **abus abuse** (12), **parricide parricide** (9), adultère adultery (8), France (9), Duhamel (8), alcoolisme *alcoholism* (7), maladie *disease* (7), silence *silence* (7), amour *love* (7), prostitution *prostitution* (6), consentement *consent* (6), **meurtre murder** (6), **guerre war** (6), mère *mother* (8), juge *judge* (6), enfance *childhood* (6), **cannibalisme cannibalism** (5), **maltraitements mistreatment** (5), **trahison treason** (5), question *question* (5), Olivier (5), relation *relationship* (5), **crime crime** (5), sodomie sodomy (4), **infanticide infanticide** (4), enfant *child* (6), homosexualité *homosexuality* (4), abandon *abandonment* (4), conséquence *consequence* (4), Isabelle (4), **suicide suicide** (4), handicap *disability* (4), mariage *marriage* (4), sujet *subject* (4), rapport *relation* (4), famille *family* (4), père *father* (4), victime *victim* (4), femme *woman* (4)

In line with the previous findings, there are a few terms referring to family in the French corpus, like *mother*, *family*, *father*, *marriage*, and *childhood*, which are absent in the English corpus. However, there are also terms denoting extreme violence, also found in the English corpus, like *murder*, *suicide*, or *cannibalism*. These terms – with the notable presence of *parricide*, always referring to the myth of Oedipus – frame incest in fiction, as shown in Table 5.10.

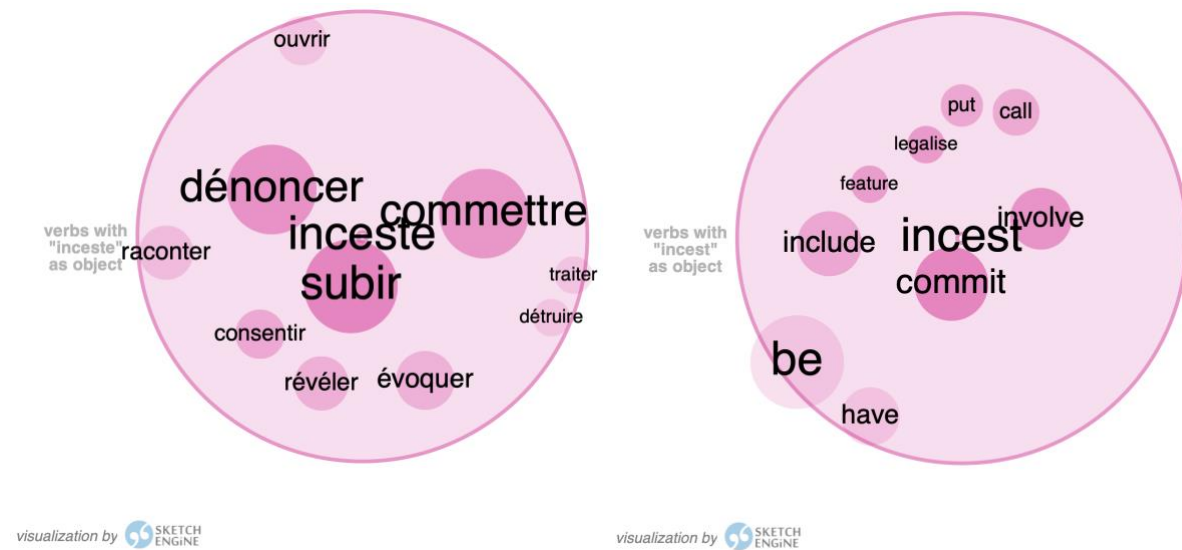
Table 5.10 Sample of concordance lines of *incest* collocating with *murder*, *cannibalism*, and *suicide* in the French corpus

| N | Np | date | co-text | node | co-text | translation |
|---|-----|------------|---|----------|--|---|
| 1 | Est | 28/03/2018 | Résumant avec humour Œdipe Roi , Œdipe à Colone et Antigone , elle a su conter les | incestes | et autres joyeusetés, pour que le public, en fin de séance, ait enfin compris ce | Summarising with humour Oedipus Rex , Oedipus at Colonus , and Antigone , she was able to recount the |

| | | | | | | |
|---|-----|------------|--|-------------------|---|---|
| | | | aventures familiales alliant meurtres , | | qu'il n'avait pas saisi jadis lors des cours d'histoire... | family adventures combining murders, incest, and other joys , so that the audience , at the end of the session, finally understood what they had not grasped in the past during history classes... |
| 2 | Fg | 18/05/2020 | Tout Shakespeare : forêts magiques, bouffons qui résumant tout en grinçant, ruses d' amour à base de métamorphoses et d'épreuves tordues, ducs mélancoliques en exil, princes usurpateurs ravagés par l'hubris, ses relations familiales qui sont bornées aux extrêmes par | l'inceste | et le meurtre , son énergie poétique , la réversibilité des sentiments sur fond d'ambivalence des passions . | All of Shakespeare: magical forests, jesters summarising everything with a grin, love intrigues based on metamorphoses and twisted trials, melancholic dukes in exile, usurping princes ravaged by hubris, his family relationships bounded to the extremes by incest and murder , his poetic energy , the reversibility of feelings against a background of ambivalence of passions . |
| 3 | Crx | 09/07/2018 | Sommet de l'horreur , l'acte ultime d'un cercle infernal de violence et de vengeance où bouillonnent vol, adultère, | inceste, | infanticide, cannibalisme. | Peak of horror , the ultimate act of an infernal circle of violence and revenge where theft, adultery, incest, infanticide, cannibalism are boiling. |
| 4 | Fg | 23/07/2018 | Imaginer assassinats, infanticides, | incestes | et cannibalisme ne soulage pourtant Sénèque qu'un temps. | Imagining assassinations, infanticides, incest, and cannibalism only provides Seneca with temporary relief. |
| 5 | O-F | 03/08/2019 | Ses textes , aux thèmes parfois durs comme | l'inceste | ou le suicide , sont renforcés par la langue de son île, le créole, et la musique traditionnelle de la Réunion, le maloya. | Her texts , sometimes dealing with harsh themes like incest or suicide , are reinforced by the language of her island, Creole, and the traditional music of Réunion, Maloya. |
| 6 | VdN | 22/01/2021 | Dans <i>My Broken Mariko</i> , manga en un volume, l' auteure explore des sujets tabous dans notre société, encore plus dans la société japonaise : le suicide , | l'inceste, | la violence et la solitude . | In <i>My Broken Mariko</i> , a one-volume manga , the author explores taboo subjects in our society, even more so in Japanese society: suicide, incest, violence, and loneliness . |

Therefore, incest is also discussed in fictional contexts in the French press. However, these similarities should be nuanced. Notably, while the English corpus often references pornography, series, or films, the French corpus tends to focus on Greek myth, theatre, music, and literature. This difference in their respective coverage choices of fiction is evident in another linguistic pattern identified with Sketch Engine. An additional functionality of Word Sketch provides insight into the verbs frequently co-occurring with *incest/e* as their object. Figure 5.6 shows the ten most frequent verbs with *incest/e* as object in both corpora.

Figure 5.6 Verbs with *incest/e* as object in the French and English corpora



To begin with, both corpora share one verb that frequently occurs in both: *commit* (14): *commettre* (31), marking incest as a crime, like in the pattern “X commits incest.” However, a notable difference emerges between the two diagrams. In the French corpus, verbs primarily refer to acts of speech: *dénoncer* (denounce) (31), *évoquer* (mention) (13), *révéler* (reveal) (11), and *raconter* (tell) (11). These verbs indicate disclosure, suggesting that incest is a subject of discussion, as in the pattern “X denounces incest”. Furthermore, other verbs imply violence: *subir* (suffer from) (34), *commettre* (commit) (31), and *détruire* (destroy) (5). It is worth noting that such verbs are absent in the English corpus. Conversely, the verbs co-occurring with *incest* as a direct object in English are *involve* (10), *feature* (4), *include* (11), and *be* (22), suggesting a pattern where incest is part of a whole: “X involves/includes/features/is incest.” These collocations are analysed to understand how incest is represented in this pattern, as illustrated in Table 5.11. The concordance lines reveal that incest is depicted as a fictional plot element, confirming the previous observation that incest tends to be fictionalised in the British press.

Table 5.11. Sample of concordance lines with *incest* as object of the verbs *include*, *involve*, and *feature* in the English corpus

| N | Np | Date | co-text | node | co-text |
|---|------|------------|--|--------------------|--|
| 1 | Mail | 15/05/2019 | [...] <i>Game of Thrones</i> has always been touted as the Playboy version of <i>Lord Of The Rings</i> . But the sex | (including incest, | rape and sexual torture) is chiefly about winning power - and when it isn't, |

| | | | | | |
|---|------|------------|--|---------------------------|--|
| | | | | | the effect is accidentally comical . |
| 2 | Sun | 01/05/2021 | Over four episodes , subjects | including incest , | stalking and murder are explored. |
| 3 | Tim | 21/09/2018 | The script then delivers a sinister biographical subplot | involving incest | and suicide . |
| 4 | Gu | 07/03/2021 | The first story featured in Midsomer Murders was <i>The Killings at Badger's Drift</i> - also the title of the first book - | involving incest | between a brother and sister , and nine deaths . |
| 5 | Tel | 30/06/2020 | He mainly specialised in comedies and dramas , but courted controversy by | featuring incest | in <i>Intimate Relations</i> (1953) and abortion in <i>The Crowded Day</i> (1954). |
| 6 | Mail | 02/01/2022 | So if you see a play such as Medea (by Euripides, describing Medea's murder of her own children); or Oedipus Rex (by Sophocles, | featuring incest , | suicide and self-blinding). |

As Table 5.11 demonstrates, the instances of incest refer to fictional contexts. As previously noted, a rich semantic field related to fiction emerges in the vicinity of the verbs, like *dragons*, *fantasy-book*, *series*, *stories*, *episode*, *comedies*, *dramas*, *play*, *version*, *episodes*, *script*, *subplot*, *story*, *title*, etc. As such, incest is a mere narrative element. Additionally, the references to *Playboy* and *Games of Thrones* in example (1) and *incest between a brother and sister* in example (4) reinforce the transgressive and eroticised representation of incest. For instance, the popular series *Games of Thrones* famously portrayed consensual incest between twins, whereas this type of incest has never been observed in anthropological research (Dussy, 2021, see Chapter 2). On a side note, it is worth pointing out that there are other instances of *incest* with *involve*, but they once again relate to abortion, where “incest is *involved*.”

In contrast with the verbs *involve*, *include*, and *feature*, the most frequent verbs with *incest* as object in the French corpus refer to acts of speech: *dénoncer* (denounce), *évoquer* (mention), *révéler* (reveal), and *raconter* (tell). While the linguistic patterns in the English corpus indicate that incest is a part of a whole, the linguistic patterns in the French corpus suggest that incest is the whole, i.e., the main topic, following the structure “X denounces/mentions/reveals/tells incest”, as seen in Table 5.12.

Table 5.12 Sample of concordance lines with *inceste* as objects of the verbs *dénoncer* (denounce), *évoquer* (mention), *révéler* (reveal), and *raconter* (tell), in the French corpus

| N | Np | date | co-text | node | co-text | translation |
|---|----|------------|---------|----------------------------|---|---|
| 1 | Mt | 13/09/2021 | Elle | dénonçait l'inceste | dont elle était victime mais à l'époque, briser le tabou était plus monstrueux que | She denounced the incest of which she was a victim , but at that time, breaking the taboo was more monstrous than the incest itself. |

| | | | | | | |
|---|-----|------------|---|-------------------------------------|--|---|
| | | | | | l' inceste lui-même. | |
| 2 | Pg | 11/03/2022 | Le 7 janvier 2021, paraissait <i>Familia grande</i> , le récit poignant de Camille Kouchner, dans lequel l' autrice | dénonce l'inceste | dont a été victime son jumeau. | On 7 January 2021, <i>Familia grande</i> was published, the poignant narrative by Camille Kouchner, in which the author denounces the incest of which her twin brother was a victim . |
| 3 | VdN | 18/09/2018 | Les paroles en créole, traduites dans le livret , sont fortes voire terribles lorsqu'elle | évoque l'inceste | dont elle a été victime . | The lyrics in Creole, translated in the booklet , are powerful and even awful when she mentions the incest of which she was a victim . |
| 4 | L'H | 12/07/2020 | Après sa mort en 1997, la publication de ses mémoires | révèle l'inceste | dont elle fut victime , qui éclairera d'un jour nouveau toute son œuvre poétique . | After her death in 1997, the publication of her memoirs reveals the incest of which she was a victim , which will shed a new light on her entire poetic work . |
| 5 | Crx | 21/09/2021 | Du point de vue des victimes , | révéler ou évoquer l'inceste | et le viol est douloureux . | From the victims' point of view, revealing or mentioning incest and rape is painful . |
| 6 | O-F | 08/11/2021 | 1986. Publie <i>Le Viol du silence</i> , un récit dans lequel elle | raconte l'inceste | qu'elle a subi à l' âge de 15 ans et le traumatisme qui s'est ensuivi. | 1986. She published <i>Le Viol du silence</i> , a narrative in which she recounts the incest she endured at the age of 15 and the ensuing trauma . |

The concordance analysis reveals that many instances of the verbs *denounce*, *reveal*, *evoke*, and *tell*, are often associated with the pain of speaking out, like the adjectives *monstrous*, *poignant*, *powerful*, *awful*, and *painful*. Furthermore, the focus is on autobiographical fiction, wherein victims of incest as sexual abuse narrate their experiences in the first-person. In doing so, incest becomes the main theme of the stories narrated by the victims themselves. As such, literature and other art forms like music serve as vectors of painful revelations. The authors disclose their abuse in a non-fiction narrative. For this reason, the translation of these linguistic patterns was challenging. Indeed, the patterns are often constructed as “*dont elle [être] victime*”, as in examples (1)-(4). It was decided to translate this literally as “of which she [be] a victim,” rather than using a more liberal translation such as “she [suffer] from,” in order to keep the keyword *victim*, which would otherwise be an omission.

Therefore, the concordance analyses of the collocates of *incest* indicate that both the French and British press mention incest within the context of fiction. However, while the British press tends to reinforce the eroticised representation of incest already widespread in contemporary fiction, the French press focuses on realistic fiction where victims narrate their experiences of incest as sexual abuse. The next section attempts to shed light on the representation of incest as intrafamilial child sexual abuse in both the French and British press.

2.2. The portrayal of incest as child sexual abuse

As discussed in Chapter 2, the definitions of incest in French and British laws differ. French law emphasises the notion of incestuous rape, while British law forbids incest when it is a consensual relationship between adults. Interestingly, British law specifies that such relationships between family members are forbidden for women “over 16 years old,” specifying the female age of consent but not for men (see Chapter 2). Furthermore, the overview of the corpora demonstrates that one of the main foci of the British press covers the abortion ban in the US, implemented “even in cases of rape or incest.” Additionally, a more thorough observation of the English corpus reveals that incest is mainly depicted in fiction, portraying transgressive but consensual relationships between family members. Conversely, despite the presence of fictional framing of incest in the French press, the primary focus is on portraying incest as child sexual abuse (henceforth CSA), as the most frequent verbs collocating with *incest* enhance the victims’ testimonies. Therefore, the comparison of the French and English corpora reveals a striking absence: the lack of representation of incest as CSA in the English corpus. This contrast sheds light on the cultural perception of the incest taboo, mirroring the legal definitions wherein France penalises incest as “rape and sexual assault”, while the UK defines it as “sexual intercourse.”

For this reason, this section examines how incest is depicted as CSA in the French press and whether such depictions exist in the British press. As such, the analysis focuses on the main social actor of CSA, namely children. Notably, the lemmas *child: enfant* frequently occur in both corpora, as mentioned previously. Firstly, the analysis centres on the verbs frequently co-occurring with *enfant* in the French corpus. Subsequently, other collocates are considered to determine if the British press covers incest as CSA. It is worth noting that in cross-linguistic analysis, each language and corpus should be analysed independently, rather than considering predetermined features. As each language has a distinct grammar, implying different linguistic patterns, the linguistic analysis may vary. For instance, there are no distinct patterns of lexical nouns modifying or modified by *enfant* because such structures are extremely rare in French, where a noun is modified by another noun with a preposition. Therefore, the analysis focuses on the most frequent verbs cooccurring with *enfant* in the French corpus, as seen in Table 5.13.

Table 5.13 The twenty most frequent verbs with *incest* as object

protéger protect (153), **maltraiter mistreat** (20), **naître be born** (24), **violer rape** (20), **placer place** (16), **être be** (118), **agresser assault** (15), **élever raise** (13), **abandonner abandon** (13), **aimer love** (14), **écouter listen** (11), **manipuler manipulate** (10), **aider help** (11), **confier entrust** (11), **laisser leave** (11), **voir see** (15), **toucher touch** (10), **entendre hear** (10), **accompagner accompany** (9), **défendre defend** (10)

The highlighted verbs are those referring to CSA, indicating the vulnerability of children. The notions of *listen* and *hear* address the need to be more attentive to the children who speak out. In order to conduct a more thorough analysis, the concordance lines of the most frequent verb, *protéger* (protect), are closely read. Out of 153 occurrences, 121 refer to incest,

underscoring the need to safeguard children who are victims of this intrafamilial abuse, as reported in Table 5.14.

Table 5.14 Sample of concordance lines of the verb *protéger* (protect) with *child* as object

| N | Np | date | co-text | node | co-text | translation |
|---|-----|------------|---|------------------------------|---|--|
| 1 | Exp | 23/01/2021 | En outre, il nous faut adapter notre droit pour mieux | protéger les enfants | victimes d’inceste et de violences sexuelles », a précisé le chef de l’État , en indiquant avoir demandé au garde des Sceaux , Éric Dupond-Moretti, et au secrétaire d’État à l’Enfance et aux Familles, Adrien Taquet, « de mener une consultation qui devra déboucher rapidement sur des propositions ». | Furthermore, we need to adapt our law to better protect children who are victims of incest and sexual violence ”, stated the president , indicating that he had asked the Minister of Justice , Éric Dupond-Moretti, and the Secretary of State for Children and Families, Adrien Taquet, “to conduct a consultation that should quickly lead to proposals”. |
| 2 | S-O | 23/06/2021 | Car, comme le rappelle la loi du 21 avril 2021 visant à protéger les mineurs des crimes et délits sexuels et de l’ inceste , et pour mieux | protéger les enfants, | aucun adulte ne peut se prévaloir du consentement sexuel d’un jeune s’il a moins de 15 ans, ou moins de 18 ans en cas d’ inceste . » | Because, as the law of April 21, 2021, aiming to protect minors from sexual crimes and offences , and from incest , reminds us, and to better protect children , no adult can rely on the sexual consent of a young person if they are under 15 years old, or under 18 years old in the case of incest . |
| 3 | DM | 27/10/2021 | La mère qui veut | protéger son enfant | du père incestueux doit être protégée, pour la Ciivise. | The mother who wants to protect her child from the incestuous father must be protected, according to Ciivise. |
| 4 | Lib | 27/10/2021 | Il faut | protéger l’enfant | lorsqu’il révèle l’ inceste , mais aussi, dans le cas où l’ agresseur est le père , protéger sa mère . | It is necessary to protect the child when they disclose incest , but also, in the case where the aggressor is the father , to protect their mother . |

Examples (1) and (2) demonstrate that the protection of children is a matter considered at the state level, with the law serving as a crucial tool to safeguard child victims of sexual abuse, and specifically incestuous abuse. Consequently, incest is presented as a public concern overseen by national authorities. Examples (3) and (4) are common depictions of incest as CSA where the figure of the father is presented as the main incest perpetrator, whereas the mother is the figure trying to protect the child. The French Commission of Inquiry against Incest and

Sexual Violence (CIIVISE), mentioned in example (3), published reports in 2021 and 2022 that shed light on the mothers' inability to take legal actions against the abusing fathers who are the main aggressors (CIIVISE, 2022, 2023, see Chapter 2). Thus, protection is the key to tackling incestuous abuse.

Additionally, the explicit term of *rape* is the fourth most frequent verb. Therefore, the verb *rape* is analysed to investigate in which context it occurs. Similarly, the findings suggest that when the verb *rape* has *child* as its object, it primarily concerns incestuous rape, with explicit mention of the filial relationship between the child and the abuser, as seen in Table 5.15.

Table 5.15 Sample of concordance lines of the verb *violier* (rape) with *child* as object

| N | Np | date | co-text | node | co-text | translation |
|---|-----|------------|--|-----------------------------|-----------------------------|--|
| 1 | O-F | 27/09/2022 | Cet hiver-là, un magicien de la région de Rennes, aujourd'hui âgé de 32 ans, avoue avoir | violé son enfant. | | That winter, a magician from the Rennes region, now 32 years old, confesses to having raped his child . |
| 2 | DL | 19/06/2020 | Au terme de quatre jours d'un lourd procès, la cour d'assises de Vaucluse, présidée par Florence Tréguier, a reconnu une Cavaillonnaise de 46 ans coupable d'avoir | violé ses enfants | et d'avoir filmé ces actes. | At the end of a four-day long trial, the criminal court of Vaucluse, presided by Florence Tréguier, found a 46-year-old woman from Cavaillon guilty of having raped her children and filmed these acts. |
| 3 | DM | 14/11/2018 | Au cœur de la Nièvre, des parents soupçonnés d'avoir | violé leurs enfants. | | In the heart of Nièvre, parents suspected of having raped their children . |
| 4 | Fg | 04/02/2019 | À l'image aussi de Céline qui, à 36 ans, accuse son père de l'avoir | violée enfant. | | Just like Céline, who at 36, accuses her father of having raped her as a child . |

Therefore, these findings align with those found in the previous subsection, confirming that the French press predominantly covers incest as CSA. There are no apparent linguistic taboos on the topic. Instead, it tends to be openly discussed in the press, explaining legal actions or various incestuous abuses that were sentenced in court. However, the British press seems to have a different coverage of the incest taboo, emphasising a more sensationalist one. For this reason, a more thorough analysis is needed. It seems doubtful that fictional incest or incest as a non-condition for having an abortion are the only ways that the British press covers the incest taboo. These topics seem predominant, but incest might be reported as CSA in the

British press as well. In doing so, the analysis actively investigates something that is missing due to extratextual knowledge of social and political contexts. This is what Duguid and Partington (2018) name *known absence*, or suspected absence, i.e., “you already know which linguistic feature you are searching for and simply want to know whether or not it is in the corpus” (p. 39, see Chapter 4). In this study, it has already been highlighted that the prevalence of incest as CSA is much higher than willingly acknowledged in society due to its uttermost tabooeness (Armstrong, 1978; Dussy, 2021; Justice & Justice, 1980; Russell, 1983, see Chapter 2). Furthermore, Partington (2014) points out that finding absence can be facilitated by comparing corpora of different genres or languages. As such, the French corpus clearly indicates that incest can be covered as CSA. Thus, the aim is to determine whether the British press portrays incest as CSA. For this purpose, the collocates of *child* reveal interesting linguistic patterns. When *child* is modified, the collocates mainly relate to abortion, but when *child* is the modifier, collocates suggest that the main topic is *child abuse* as Tables 5.16 and 5.17 show respectively. This demonstrates that the British press does cover child abuse.

Table 5.16 The modifiers of *child* in the English corpus

unborn (138), **young** (50), migrant (13), **eldest** (12), disabled (10), only (18), **preborn** (7), unaccompanied (7), cellar (7), **many** (24), **more** (22), **own** (20), small (10), **illegitimate** (6), innocent (6), vulnerable (6), dead (6), **fourth** (6), other (22), adult (5), real (8), transgender (5), secret (5), love (5), **stillborn** (4), citizen (4), old (7), adult (4), **biological** (4), **third** (5), naked (4), school (5), **first** (14), **second** (6), **future** (4), poor (4), woman (4), few (4), US (4)

Table 5.17 Nouns modified by *child* in the English corpus

abuse (107), **pornography** (12), **material** (11), **victim** (15), support (12), poverty (7), **image** (10), **molestation** (5), **molester** (4), sacrifice (6), care (14), charity (5), marriage (7), abuser (4), **rapist** (4), **rape** (5), star (5), **video** (4), **expert** (4), actor (4), protection (4), **case** (4), law (6)

From the list of collocates with the modifiers of *child*, the main focus is on abortion, with mention of numbers or the age of children, as highlighted in Table 5.16. However, when *child* is the modifier, as seen in Table 5.17, *child abuse* is a frequent collocation in the British press. It remains to be determined whether they refer to incest, as each article was collected in the corpus through the search term *incest*. From this list of collocates with *child* as a modifier, two observations can be formulated: (i) with the exceptions of *star* and *actor*, all the nouns modified by *child* refer to vulnerability, including *poverty*, *sacrifice*, or *protection*, (ii) most of the terms refer to child **sexual** abuse, including terms like *expert* or *case*, as in “child abuse expert” or “child rape cases,” and are thus highlighted. However, there is no mention of incest in particular. For this reason, the concordance lines of *child abuse* are analysed to determine in which context this term is used, and whether it may refer to incest. After closely reading the 107 concordance lines of *child (sex*) abuse*, it was found that 26 refer to incest, either explicitly by using the term *incest* in the same sentence, or implicitly by mentioning the family environment where abuse occurs. Table 5. 18. provides a few examples of such occurrences.

Table 5.18 Sample of concordance lines of the collocation *child abuse* implying incest

| N | Np | date | co-text | node | co-text |
|---|-----|------------|---|---------------------------|--|
| 1 | Gu | 18/01/2018 | In October the high court held that Victorian courts had been handing down inadequate sentences for one category of crime - | child abuse | in families - owing to a 1968 precedent in an incest case. |
| 2 | Sun | 30/05/2018 | Barr, who at 16 spent eight months in a mental institution after suffering anxiety and hallucinations following a near-death car accident, told a large congregation at a church in Denver in 1991 that she was a victim of incest and | child abuse. | |
| 3 | Ind | 03/02/2020 | And yet, around two-thirds of | child sexual abuse | takes place within the family environment , according to an inquiry by England's children's commissioner in 2015. |
| 4 | Tim | 11/02/2021 | Addressing the victims of incest and | child sex abuse | after the Duhamel scandal erupted, President Macron said that the national awakening meant that shame had changed sides. |

As such, the explicit mention of incest related to *child (sex*) abuse* only represents a small percentage of the concordance lines. Most of the other concordance lines use *child abuse* as a euphemism for *paedophilia*, as seen in examples (1) and (2) in Table 5.19, or *child pornography*, as illustrated in examples (3) and (4). Indeed, it appears that *child abuse* can serve as a euphemistic alternative to *paedophilia* or *pornography*. Nevertheless, the terms *paedophilia/pedophilia* and *pornography* occur in the corpus 76 times and 146 times respectively. Thus, using *child abuse* instead of these terms is not necessarily a sign of avoidance but rather an alternative for journalists to avoid repetitions. However, it is worth noting that *child abuse* encompasses two notions, *paedophilia* and *pornography*, making it vague and more ambiguous. As such, it can be considered euphemistic due to this ambiguity (see Chapter 3).

Table 5.19 Sample of concordance lines of the collocation *child abuse*

| N | Np | date | co-text | node | co-text |
|---|-----|------------|---|----------------------------|--|
| 1 | Gu | 25/05/2018 | The vote was a reminder, she said, of the church's loosening grip on a country where a series of scandals , involving | child abuse | and mistreatment of pregnant, unmarried women and their children, have hugely undermined the clergy's authority. |
| 2 | Tel | 27/07/2021 | Headlined "Radical thoughts on consent", Mr Tatchell wrote that while the authors | child sexual abuse" | they "also argue that consenting, victimless sexual relationships between younger |

| | | | | | |
|---|------|------------|---|--------------------|--|
| | | | "oppose coercive and exploitative | | and older people should not be penalised by the law". |
| 3 | Mail | 11/12/2019 | A married vicar who was caught with nearly 2,000 | child abuse | images has been jailed for 20 months. |
| 4 | Gu | 25/03/2020 | MPs and campaigners are calling for urgent action to stop videos of rape, revenge porn and | child abuse | being posted on Pornhub as traffic to the site booms amid a worldwide Covid-19 lockdown. |

Therefore, *child (sex*) abuse* can directly or indirectly refer to *incest* but may also refer to *paedophilia* or *pornography*. Nevertheless, analysing the collocates of *child* and then the concordance lines of *child abuse* made it possible to find the known absence: the British press does cover incest as CSA, even though it is not its predominant focus. Thus, the known absence became a known presence, albeit a fragile one. Consequently, the aim is to further analyse the British coverage of incest as CSA to determine whether the corpus linguistic measures have overlooked any elements so far, as corpus linguistic tools may be overarching and miss details (Baker, 2023). For this purpose, the concordances of the four collocates of *child* as modifier indicating social actors, i.e., *victim*, *molester*, *abuser*, and *rapist*, are analysed. The first term examined is *victim*, occurring 15 times with *child* as modifier. Over the 15 concordance lines, 7 refer to incest. However, only three occurrences directly refer to incest, as seen in examples (1) and (2) in Table 5.20. The other concordance lines concern abortion, as seen in examples (3) and (4). Once again, the child victim of incest is mainly framed within the topic of abortion. Hence, the focus is not on the victims of incest, but rather on the victims of incest facing abortion. The other eight occurrences concern children who are victims of sexual violence in general, without specific mention of incest, and mainly regarding their access to abortion.

Table 5.20 Sample of concordance lines of *child victim* with *incest* as a close collocate

| N | Np | date | co-text | node | co-text |
|---|-----|------------|---|-------------------------------|---|
| 1 | Sun | 14/07/2018 | Two | child victims | of incest and neglect have launched a £238million personal injury claim, the highest in British history. |
| 2 | Gu | 13/03/2021 | But also central to the formula of incest abuse are the denial and disbelief of those around the | child victim. | |
| 3 | Tim | 22/05/2018 | "Everything should be in place to allow | child victims | of rape and incest , who are in a position, the right to choose ," she said. |
| 4 | Ind | 13/10/2022 | | Child victim of incest | was denied abortion in Florida. |

Thus, the examples illustrate how child victims are represented as powerless and resourceless. As Jewkes (2015) observes, the child victim is a newsworthy item, especially when it is linked to crime. Furthermore, the child aggressors are also depicted in the press, as the list of

collocates indicates (Table 5.17). The concordance lines of child *molester*, *abuser*, and *rapist* are analysed to understand their portrayal, and whether they are related to incestuous abuse. However, once again, the analysis demonstrates that the references to incest are rare. The representations of child abusers refer to different contexts, but mainly to paedophilia. For instance, example (1) in Table 5.21 shows how the journalist links the large category of “child abusers” with an Afghan veteran by reporting Trump’s opinion on the death penalty. The association of both demonstrates a confusion where two very different contexts are linked together through the same condemnation: the death penalty. It does indicate that child abuse is strongly prohibited and penalised officially. Similarly, the reported speech in example (2), claiming to take child rapists “off the streets,” is a common statement to assert that actions are undertaken against this type of criminals. However, the mention “off the streets” contributes to the rape myth consisting in blaming unknown criminals lurking in dark streets at night, instead of acknowledging that the large majority of rapes are perpetrated by relatives (Brownmiller, 1993; see Chapter 2). Indeed, “the figure of the paedophile reaches a consensus as the ‘identifiable hate-figure’” who is “nearly reduced to a set of sub-humans, even bestial” (Jewkes, 2015, pp. 118–119). Thus, it is easier to blame an evil category rather than to recognise structural problems contributing to CSA, such as incestuous abuse within the families.

In fine, there are only two lines of all these collocates that refer to incest perpetrators, but even in these cases, it is implicit. Example (3) indirectly mentions the accusation of incestuous rape against Woody Allen made by his adoptive daughter (see Chapter 6 for further analysis). Example (4) is the testimony of the victim who uses the term “child molester” without specifying to whom it refers. However, the intimacy of the abuse, happening in the house, “if you grew up in a home,” and enhancing the inaction of the mother, “with a mother who accepts it and enables it,” correspond to anthropological descriptions of incestuous abuse (Armstrong, 1978; Dussy, 2021; Justice & Justice, 1980). By taking a closer look at this article published in the *Telegraph* in November 2021, it is understood that incest is indeed mentioned and explained at length. However, the article is about the American murder case of the “The Menendez brothers” who shot down both their parents. They testified that their father molested them from the age of six, and that the “fondling” “incrementally progressed up to rape.” However, the fact that one of the rare articles mentioning incest involves cold-blooded murder demonstrates that the sensationalism around incest is acute. Incest is discussed because it involves parricide and matricide, which could be reminiscent of the Oedipus myth to a certain extent. Even though the brothers explain the mechanisms of incest they suffered from, such explanations are obfuscated because the case remains extraordinary, in the sense that it is rare. Not all the victims of incest murder their parents. Therefore, one of the rare articles mentioning the mechanisms of incestuous abuse contributes to representing incest with extreme violence, in the same way as it is described in fiction.

Table 5.21 The concordance lines of *child molester*, *child abuser*, and *child rapist*, in the English corpus

| N | Np | date | co-text | node | co-text |
|---|------|------------|--|-----------------------------|---|
| 1 | Mail | 17/12/2022 | Trump rarely mentioned capital punishment during his 2016 election campaign, but he has since called for the death penalty for | child molesters | and even Afghan veteran Bowe Bergdahl, who was captured by the Taliban in 2009 after deserting his post and held for five years before being released as part of a prisoner exchange. |
| 2 | Gu | 13/07/2022 | Yost issued a single sentence statement: "We rejoice anytime a | child rapist | is taken off the streets. " |
| 3 | Gu | 13/03/2021 | Most accused | child sexual abusers | are not famous movie directors, as Allen is. |
| 4 | Tel | 27/11/2021 | "But it's not that surprising, if you grew up in a home with a | child molester, | and you grew up with a mother who accepts and enables it, and there's a lot of brutality ". |

The English corpus is therefore challenging in the sense that it raises the question whether known absence has reachable textual evidence. It calls into question whether the British press scarcely addresses incest as CSA, or whether the CL tools are too broad, and thus overlook the known absence. To explore whether further investigation might yield different insights, a different angle is chosen, consisting in examining the frequency lists in a different light. The most frequent adjectives in the English corpus are observed, and interestingly, the adjective *sexual* ranks in the top ten, as shown in Table 5.22. Alongside *legal*, they are the first two adjectives that relate to the topic of incest.

Table 5.22 The ten most frequent adjectives in the English corpus

more (4,028), other (3,616), new (3,553), last (3,159), first (3,052), many (2,806), good (2,610), own (1,961), legal (1,860), **sexual** (1,793)

The fact that *sexual* is one of the most frequent adjectives is consistent with the corpus content. Thus, the analysis focuses on the adjective *sexual* to explore its potential contribution to depicting incest as CSA. As seen in Table 5.23, the main collocates of *sexual* as a modifier have negative semantic prosody (made in bold) or positive one (italicised). The terms *survivor* and *victim* are both made in bold and italicised because it depends on the way they are perceived. Indeed, being a *sexual survivor* or *sexual victim* means that the person underwent sexual violence, but also overcame it. They have both negative and positive meanings.

Table 5.23 The collocates of *sexual* as modifier

assault (255), **abuse** (186), violence (185), **harassment** (96), *relationship* (54), **misconduct** (38), *orientation* (29), **survivor** (25), **allegation** (24), *relation* (23), *act* (23), **victim** (23), **crime** (18), *activity* (16), content (14), **exploitation** (13), *freedom* (14), *consent* (12), *encounter* (12), *attraction* (12), *revolution* (12), *desire* (11), *partner* (11), **predator** (10), advance (10), penetration (9), *intercourse* (9), **tension** (9), behaviour (9), case (15), politics (9), contact (8),

image (9), **scandal** (8), **perversion** (7), gratification (7), **offence** (7), *interest* (7), *liberation* (6), *awakening* (6), *fantasy* (6), *identity* (6)

Interestingly, from all the various collocates, the term *relationship* yields the most significant findings. It was initially perceived as positive but revealed as negative through concordance analysis. Out of 54 occurrences, 34 concordances refer to incestuous relationships, with a predominant linguistic pattern: “sexual relationship + [preposition WITH/BY] + family member.” Unexpectedly, the collocation *sexual relationship* is the most relevant one to uncover the known absence of incest as CSA, as illustrated in Table 5.24. These findings suggest a tendency to employ linguistic avoidance strategies to euphemise incest. However, despite clear references to kinship in sentences, ambivalence persists in examples (1), (2), and (3). Words like “embark,” “initiate,” and “consensual” indicate consent, but the mentions of time and age are crucial. The transgression of the incest taboo is mitigated either by specifying that the family members met after a long time, implying that they barely know each other, or that the stepdaughter was “over the age of 18,” referring to the British law stating that women “over the age of 16” should “not permit” the sexual intercourse, rendering it illegal. These examples illustrate that sexual relationships with a family member give rise to discomfort, expressed through condoning elements. To some extent, the condoning factors are mentioned precisely to mitigate the discomfort, like the rhetorical question in example (4). Finally, examples (5) and (6) imply incest as CSA where abuse is inferred from the mention of disagreement with the verbs “deny” and “embroil”.

Table 5.24 Sample of concordance lines with the collocation *sexual relationship*

| N | Np | date | co-text | node | co-text |
|---|------|------------|---|-----------------------------|---|
| 1 | Mail | 30/07/2019 | Her father , pianist Joaquin Nin, reappeared in her life after a 20-year absence and they embarked on a | sexual relationship. | |
| 2 | Tim | 11/01/2022 | She argued that a stepfather who had a consensual | sexual relationship | with a stepdaughter over the age of 18 should not face prosecution. |
| 3 | Ind | 11/02/2022 | Authorities have said that the siblings met and allegedly initiated a | sexual relationship. | |
| 4 | Sun | 13/03/2021 | “You do know already that having any romantic or | sexual relationship | with a sibling is not allowed”. |
| 5 | Mir | 17/07/2022 | Ricky Martin has strongly denied claims by his nephew that they had a | sexual relationship. | |
| 6 | Ind | 07/09/2022 | Dr Oz got embroiled in the subject when he was asked on the show about advising a listener involved in a | sexual relationship | with their cousin. |

Thus, the analyses appear to reach a relevant dead-end, which is due to the scarcity and particularity of the representation of incest as CSA in the British press. However, its relevance lies in confirming the topic's tabooess, evident in its absence from press coverage. Indeed, there is a discrepancy between what is reported in the literary review and what is represented in the press. Even though the French press favours another approach, the British press seems to illustrate the success of the taboo, meaning that a taboo is truly successful when it is not only euphemised but avoided altogether, thus falling into silence (Popescu, 2018a). Thus, the fact that incest is scarcely reported as CSA in the British press suggests that the taboo is too strong to have explicit textual evidence. As a consequence, this study may reach the limits of finding discursive absence. However, the aim of corpus-assisted discourse analysis is to highlight what seems hidden in the text in the first place, revealing what is not consciously perceptible (Baker, 2023; Partington et al., 2013). Drawing on the legal definitions of incest in both countries, the cross-linguistic comparison demonstrates that newsworthiness is politically determined. Thus, the analysis highlights how the representation of incest culturally differs.

2.3. Metaphors and paradox of the adjective *incestuous*

While the previous subsection has focused on the term *incest*, highlighting the depiction of incest as a consensual relationship in the British press and as CSA in the French press, this subsection aims to explore the usage of the adjectives *incestuous: incestueux*. The adjective is less frequent in the English corpus, with 486 occurrences, than in the French corpus, where it occurs 1,792 times. This indicates that the adjective *incestuous: incestueux* is used three times as frequently in the French corpus than in the English one when relative frequencies are considered.

As expected, and in line with the previous findings, the adjective *incestueux* is used more literally in the French corpus, pertaining to CSA, with terms indicating male family members like *father, grandfather, uncle, stepfather, parent, and brother*, as seen in Table 5.25. Conversely, the adjective *incestuous* primarily qualifies consensual incest in the English corpus, with words like *affair, desire, sex, scene, passion, marriage, porn, and fantasy*, as shown in Table 5.25.

Table 5.25 The twenty most frequent nouns modified by *incestueux: incestuous* in both corpora

| N | French corpus | Translation | English corpus |
|---|-------------------|--------------------|----------------|
| 1 | viol | rape | relationship |
| 2 | <u>père</u> | <u>father</u> | <i>affair</i> |
| 3 | relation | relationship | union |
| 4 | acte | act | <i>desire</i> |
| 5 | <u>grand-père</u> | <u>grandfather</u> | world |
| 6 | caractère | nature | abuse |
| 7 | <u>oncle</u> | <u>uncle</u> | relation |
| 8 | <u>beau-père</u> | <u>stepfather</u> | sex |

| | | | |
|----|--------------|-----------|-----------|
| 9 | lien | tie | sibling |
| 10 | geste | gesture | scene |
| 11 | agression | assault | passion |
| 12 | climat | climate | rape |
| 13 | amour | love | marriage |
| 14 | parent | parent | undertone |
| 15 | rapport | relation | mother |
| 16 | comportement | behaviour | porn |
| 17 | attouchement | molesting | circle |
| 18 | fait | fact | elite |
| 19 | frère | brother | fantasy |
| 20 | agissement | doing | nature |

However, by comparing the use of *incestuous* with *relationship: relation* both in English and French, one unexpected finding came across the analysis: the collocations *incestuous relationship: relation incestueuse* are sometimes used metaphorically in both languages. In this sense, it refers to power collusion. As Table 5.26 shows from the English corpus and Table 5.27. from the French corpus, journalists use this term figuratively to criticise what are considered unhealthy relations between establishment stakeholders.

Table 5.26 Sample of concordance lines with the collocation *incestuous relationship* used metaphorically in the English corpus

| N | Np | date | co-text | node | co-text |
|---|-----|------------|--|--------------------------------|--|
| 1 | Ind | 16/09/2018 | On top of dismay about Blair's wars, there was creeping disquiet about New Labour's | incestuous relationship | with bankers and neglect of Britain's decaying regions. |
| 2 | Gu | 07/09/2019 | Trump's has long had a tempestuous and | incestuous relationship | with the media. |
| 3 | Tim | 20/08/2018 | Antonio Polito, a commentator writing in <i>Corriere della Sera</i> , said that the populists had convinced the public that an | incestuous relationship | between the state and Autostrade meant that it was not supervised adequately. |
| 4 | Mir | 06/04/2022 | So you would like to know about the | incestuous relationship | between politicians and journalists? |

Table 5.27 Sample of concordance lines with the collocation *relation incestueuse* used metaphorically in the French corpus

| N | Np | date | co-text | node | co-text | translation |
|---|----|------------|---|-------------------------------|--|--|
| 1 | Fg | 05/12/2017 | Mauvaise gestion, gaspillage, programmes et contenus médiocres, | relations incestueuses | entre l'audiovisuel et ses partenaires tels que les animateurs ou les producteurs... rien ou presque n'a été épargné. | Mismanagement, waste, mediocre programmes and contents, incestuous relationships between broadcasting and its partners such as hosts and producers... |

| | | | | | | |
|---|-----|------------|--|---------------------------------------|---|--|
| | | | | | | nothings or almost nothing has been spared. |
| 2 | L'H | 19/12/2017 | Mais il est un grief qu'a formulé le président que nous partageons , celui des | relations incestueuses | avec les producteurs privés « abonnés à la commande publique ». | But we share one grievance that the president has expressed, that of incestuous relationships with private producers who are “subscribers to public contracts”. |
| 3 | O-F | 19/01/2020 | Caen où Nicolas Bay dénonce | « les relations incestueuses » | qu'entretiennent Les Républicains et La République en marche. | It's in Caen where Nicolas Bay denounces the “ incestuous relationships ” between the conservative party and Macron's party. |
| 4 | Ech | 15/05/2020 | Et que dire des | relations incestueuses | entre certains potentats du marché de l'art et les institutions publiques. | Let alone the incestuous relationships between certain tycoons of the art market and public institutions. |

Interestingly, both languages display similar linguistic patterns denoting negative semantic prosody: *incestuous relationships* are harshly condemned with co-occurring words inferring inefficiency, such as *dismay, disquiet, neglect, mismanagement, waste, grievance, decaying, tempestuous, mediocre, not adequately, nothing spared, denounce, populist, and tycoon*. Thus, collusions between political power and other institutions, like the media, banks, and private stakeholders, are brought to the public eye. The metaphor lies in the fact that these influential entities are not involved in sexual relationships but maintain an unhealthy closeness between them. Metaphors are conceptual linguistic devices that transfer the notion of a source domain (e.g., incest in the family) into a target domain (e.g., network of power) (Lakoff & Johnson, 1980). The metaphorical use of the adjective *incestuous* serves to disqualify this type of network, where everyone seems to know each other. That is why other collocates of *incestuous* include the words *circle, elite, world, clique, and union*, as seen in Table 5.25. These terms serve to condemn the relationships between prominent figures, implying that they belong to the same family. Sometimes, adverbs such as *somewhat, quite, almost, slightly*, or the use of brackets, like in example (3) in Table 5.28, reinforce the metaphorical use of *incestuous* by enhancing the journalist's uneasiness in properly characterising these interconnected relations.

Table 5.28 Sample of concordance lines with the adjective *incestuous* used metaphorically in the English corpus

| N | Np | date | co-text | node | co-text |
|---|-----|------------|---|--------------------------|---|
| 1 | Sun | 09/11/2018 | And in truth, this odd decision has been thrashed out between a small, | incestuous circle | of PR men and agents, with Southgate presumably shrugging his shoulders and deciding that it was not worth fighting against. |
| 2 | Gu | 12/12/2018 | Most of these speakers were brought in through | incestuous elite. | |

| | | | | | |
|---|------|------------|---|-------------------------|---|
| | | | personal connections among Britain's | | |
| 3 | Mail | 13/05/2019 | In the somewhat | incestuous world | of cricket journalism , Agnew has been ruffling feathers for some time. |
| 4 | Sun | 18/05/2020 | And the | incestuous union | between the Premier League and Sky Sports has altered the face of English football beyond recognition. |

Likewise, the words *relation* (relationship), *rapport* (relation), *lien* (tie), and *liaison* (affair) collocate with *incestuous* with a figurative meaning in the French corpus, as shown in Table 5.29.

Table 5.29 Sample of concordance lines with the adjective *incestueux* used metaphorically in the French corpus

| N | Np | date | co-text | node | co-text | translation |
|---|-----|------------|---|------------------------------|---|---|
| 1 | L'H | 27/04/2018 | Le corps politique entretient avec le spectre médiatique des | rapports incestueux. | | The body politic maintains incestuous relations with the media spectre. |
| 2 | Fg | 01/09/2020 | Normalement passés sous silence, ces | liens incestueux | entre la politique et la justice libanaise ne font qu'aggraver le climat de défiance dans le pays. | Usually kept silent, these incestuous ties between politics and Lebanese justice only worsen the climate of mistrust in the country. |
| 3 | Fg | 23/09/2019 | Il ne fait aucun doute que la loi Bertrand du 29 décembre 2011, relative au renforcement de la sécurité sanitaire du médicament et des produits de santé, a mis fin à nombre de | liaisons incestueuses | entre l'agence et les industriels du médicament. | It's certain that the Bertrand law of December 29, 2011, concerning the reinforcement of the health safety of medication and health products, put an end to many incestuous affairs between the agency and pharmaceutical manufacturers. |

Therefore, it is worth noting that the figurative use of *incestuous* serves the same function in both languages: they denote disapproved collusions between circles of power. As such, these articles are found in the corpora because they were retrieved through the lemma *incest**. Their general topic is about politics, and not about incest as such. That is why the analysis is interesting because the adjective *incestuous* is both metaphorical and paradoxical. Indeed,

the paradox lies in the fact that the figurative meaning converges in both languages but diverges with the literal meaning. By examining the other occurrences of *incestuous relationship: relation incestueuse* when used literally, the previous findings are confirmed. While the French corpus uses the collocation of incestuous relationship to depict CSA, the English corpus primarily uses it to depict fictional or sensational incest. Table 5.30 provides a sample of such occurrences in the British press. In example (1), there is an overlap between fiction and reality with the historical figure of King Antiochus being staged on scene. Example (2) is another case of fictionalised and eroticised incest, with words like *episode, novel, feature*, as well as *bare bottom* and *sex act*. The concordance analysis reveals that the instances of *incestuous relationship* are embedded in a fictionalised representation of incest or murder. Examples (3) and (4), the coverage of incest involves murder, with terms like *evil* and *twisted*, where it is either the parents who murdered their children (3) or the children who murdered their parents (4).

Table 5.30 Sample of concordance lines of *incestuous relationship* used literally in the British press

| N | Np | date | co-text | node | co-text |
|---|------|------------|---|--------------------------|--|
| 1 | Ind | 10/04/2018 | In the opening episode, our wife-seeking protagonist solves a deadly riddle that betrays the secret of King Antiochus's | incestuous relationship | with his daughter. |
| 2 | Tim | 27/07/2019 | The first episode of Sanditon, based on her unfinished final novel, features three men's bare bottoms, the hint of an | incestuous relationship | and a sex act in the woods of a country house. |
| 3 | Mail | 13/11/2019 | An evil mum and the twisted half brother she was in an | incestuous relationship | with were jailed for life yesterday for murdering two of their children. |
| 4 | Gu | 28/08/2020 | A serious case review into the murder of two children by their parents, who were in an | incestuous relationship, | has concluded there is no evidence that agencies and professionals could have predicted such events. |

Conversely, the French press primarily uses the term *relation incestueuse* to describe CSA, as seen in Table 5.31. Examples (1) and (2) illustrate the duration of incestuous abuse, with the mentions of the female victims' age, and their filiation to their father. Example (3) reports Duhamel's defensive speech following the revelations of his rapes committed against his stepson in the book *La Familia grande*, published by his stepdaughter (Kouchner, 2021). The journalist undermines the reported speech by referring to him as "our 'great intellectual'" within brackets. Furthermore, Duhamel's arguments are characterised by phrases like *ambiguous remarks* and *tending to affirm*. Regarding the term "reciprocity" enclosed in brackets, it remains unclear whether it is bracketed as a quoted word, or whether the journalist intends to create distance from it, or both.

Table 5.31 Sample of concordance lines of *relation incestueuse* used literally in the French press

| N | Np | date | co-text | node | co-text | translation |
|---|------|------------|--|-----------------------|--|--|
| 1 | O-F | 09/06/2018 | De juillet 2015 à avril 2016, il avait entretenu une | relation incestueuse | avec sa fille, alors âgée de 14 ans, dans la maison familiale, près de Rennes. | From July 2015 to April 2016, he had maintained an incestuous relationship with his daughter, who was then 14 years old, in the family home in Rennes. |
| 2 | Pg | 20/12/2019 | À 17 ans, cette jeune fille a eu le courage de lever le voile sur la | relation incestueuse | dans laquelle son père l'avait enfermée depuis ses 9-10 ans. | At 17 years old, this young girl had the courage to unveil the truth about the incestuous relationship in which her father had confined her since she was 9-10 years old. |
| 3 | L'Hu | 18/02/2021 | En revanche, lorsque le beau-fils d'Olivier Duhamel fut violé, notre « grand intellectuel » se répandit en propos ambigus, tendant à affirmer qu'à ce même âge on était souvent conscient et qu'une forme de « réciprocité » pouvait exister dans la | relation incestueuse. | | However, when Oliver Duhamel's stepson was raped, our "great intellectual" made ambiguous remarks, tending to affirm that at the same age one was often conscious and that a form of "reciprocity" could exist in the incestuous relationship. |

Therefore, the adjective *incestuous* can be used figuratively and literally, which has not been observed for the noun *incest*, only used literally. Interestingly, the figurative meaning serves the same function in both languages, namely the denunciation of power collusions. Thus, it is paradoxical to observe that when incest is understood metaphorically, both languages converge in the definition. It is paradoxical because the source domain of the metaphors should refer to the same definition of incest in both languages in order to transfer the same mapping to the target domain, which is not the case. When incest is used literally, there are cross-cultural variations in the definition of incest, with an emphasis on CSA in the French press and on consensual relationships in the British press. Consequently, the meaning is the same on a conceptual level, where the understanding of what is incest is abstract. However, the meaning differs when the perception of incest is more specific.

Finally, this section began with an exhibition on sins at the National Gallery in London, and concludes with another exhibition staged at the British Museum, titled *Nero: The Man Behind*

the Myth, as reported in an article by *The Telegraph* published in May 2021. This exhibition aims to shed light on the real Nero, whose reputation was tarnished by his rivals, and to deconstruct myths surrounding him, such as his incestuous relationship with his mother, his matricide, his second wife's murder, and his singing during Rome's burning. However, the journalist poses the following question:

But are these stories true? [...] It is next to impossible now to sort the facts about Nero from the fiction. Some of what we read is probably outlandish fantasy. [...] The British Museum exhibition takes another approach. It is not an exercise in rehabilitation. But it uses the art, archaeology and material culture to offer different perspectives on Nero.

Similarly, many historical figures are said to have been involved in consensual incestuous relationships. However, we can wonder whether such descriptions also contribute to the creation of dark legends surrounding them. Nero, a flamboyant character who engaged in an incestuous relationship with his mother before killing her, is reminiscent of similar press articles analysed above. Therefore, taboos reveal their ambivalence in our societies, where their transgression is appealing, as noted by Freud (1913). This is why the next section examines how the concept of taboo is commented on in the press.

3. The metalinguistics of the incest taboo

Chapter 3 introduced the social and linguistic manifestations of taboo, especially those related to sexuality. In addition, it was observed that politicians and journalists often describe a topic as taboo through metalinguistic comments such as “this is a taboo,” or emphasise the boldness of breaking a taboo with phrases like “there is no taboo,” implying that, actually, there is one (Pražuch, 2018). The media, as a mirror of societies, offer introspection into what is considered taboo in cultures, such as the French and British cultures. Hence, this final section focuses on the very notion of taboo within the two datasets under scrutiny. The term *taboo* occurs 313 times in the French corpus and 102 times in the English corpus. In terms of relative frequency, the term *taboo* is 4.6 times more frequent in the French corpus than in the English corpus. Furthermore, the noticeable absence of the incest taboo as CSA has already been pointed out regarding the British press. Thus, the following analysis aims to reveal what is considered taboo in the British press, and such findings are contrasted with those from the French press.

3.1. The ambivalence of the incest taboo in the British press

As mentioned above, the term *taboo* occurs 102 times in the English corpus. Two main contexts emerge from the concordance analysis: incest and abortion. The incest taboo is referred to in 42 concordances, i.e., one-third of the occurrences of *taboo*, whereas abortion is referred to in 21 concordances, i.e., one-fifth of the occurrences. It means that despite the prevalence of abortion coverage in the English corpus, the term *taboo* primarily pertains to incest. However, there are mentions of other taboos, some of which show confusion and

moral judgments, particularly in reference to topics like homosexuality. Finally, the term *taboo* can be used to comment on the taboo, i.e., taboo is used metalinguistically.

First, the analysis focuses on the positive and negative prosodies of *taboo* in reference to incest. Interestingly, some concordance lines depict incest as CSA, as seen in Table 5.32. In this case, the incest taboo is framed by a negative semantic prosody. Example (2) reports the high-profile scandal in France in 2021. Examples (1) and (3) show similar patterns, portraying incest as CSA that “*remain[s]* taboo” or “*remained* largely in the dark”. This notion of “remaining,” i.e., staying as it is, is also characterised in example (2) with the mention of “*longstanding* social taboo.” The mentions of speaking out, like “*broken*” in example (2) or “*brought out of the shadows*” in example (3) highlight the challenges faced by victims in voicing their experiences. Here, the incest taboo is described at the intersection of social and linguistic taboos. Silence characterises the absent discourse on incest as CSA in the English corpus. In other words, the examples illustrate how incest is described as a taboo when it is depicted as CSA. Thus, these findings acknowledge the silenced presence of incest as CSA in the British press.

Table 5.32 Sample of concordance lines with negative semantic prosody characterising incest as a taboo

| N | Np | Date | co-text | node | co-text |
|---|------|------------|---|---------------|---|
| 1 | Gu | 13/03/2021 | But in spite of this disturbing prevalence , public discussions of incestuous child abuse remain | taboo. | |
| 2 | Tim | 11/01/2022 | The longstanding social | taboo | over the subject was, however, broken last year with the publication of a bestselling book claiming Olivier Duhamel, 71, a well-known political pundit, had sexually abused his teenage stepson in the 1980s. |
| 3 | Mail | 30/04/2022 | Where other forms of sexual violence have been brought out of the shadows, survivors encouraged to share their experiences in the pursuit of justice and healing, the | taboo | around incest is so powerful it has remained largely in the dark . |

Nevertheless, such descriptions of the incest taboo as CSA are rare. Most of the times, the word *taboo* in reference to incest is used in an eroticised way, where the transgression of the taboo is appealing, as shown in Table 5.33. Phrases like “*titled*,” “*to toy*,” “*depraved* incest and rape *fantasies*,” “*fantasising* about *taboo* topics,” “*turned on*,” “*aroused* by *forbidden* taboo relationships,” and “*at some level desire* it,” clearly demonstrate that the incest is appealing due to its tabooess. The eroticism resides in the transgression. As such, language

is ambiguous as eroticism is enhanced in the transgression with term like “delicious” to describe the incest taboo in example (1).

Table 5.33 Sample of concordance lines with positive semantic prosody characterising incest as a taboo

| N | Np | Date | co-text | node | co-text |
|---|-----|------------|---|----------------|--|
| 1 | Tim | 21/11/2017 | You feel that Rossini was more titillated than deeply moved by the story's possibilities, and especially by being able to toy with the delicious | taboo | of depicting incest on stage (only toy with it, mind; it took Wagner to go the full Monty). |
| 2 | DM | 01/02/2018 | Former North Lanarkshire councillor Fagan shared depraved rape and incest fantasies on | "taboo" | websites in 2016. |
| 3 | Sun | 02/02/2018 | Disgraced Fagan insisted he was just "fantasising" about | "taboo" | subjects . The confessed porn addict claimed: “I developed this fantasy round having sex with a parent and daughter. ” |
| 4 | Gu | 05/02/2018 | “I have found that many of the people that are turned on by incest pornography or fauxcest pornography are aroused by forbidden, | taboo | relationships,” she says. |
| 5 | Gu | 24/08/2018 | Still, it is said that incest wouldn't need to be | taboo | if we didn't also at some level desire it. |

Thus, it tends to corroborate Freud’s (1913) analysis that the ambivalence of the taboo lies in the desire to transgress it. This ambivalence becomes evident concerning the incest taboo insofar as language is ambiguous in portraying it as a desired transgression. To some extent, we can wonder whether this ambivalence signifies hypocrisy, as incest is described as a taboo primarily to emphasise its sensationalism.

In addition, Table 5.34 illustrates the various explanations or theories on the incest taboo. As discussed in Chapter 2, certain theories lack scientific validity, such as the syndrome of the parental alienation developed by the psychiatrist Richard Gardner in the 1980s. Similarly, example (1) mentions the hypothesis of the Genetic Sexual Attraction (GSA), which initially describes “erotic feelings between close relatives, often between siblings or between parents and children, who are separated early in life and reunited in adolescence or adulthood” (Colman, 2015). This concept was initially used in the context of adoption when biological parents and children, separated for years, reconnect, and feel sexually attracted to each other. Despite its presence in documents on adoption (Cumbria County Council), this theory is regarded as pseudo-science due to its lack of scientific foundation (Marcotte, 2016). Interestingly, example (1) suggests that siblings may feel sexually attracted to each other because they are “not protected by the *normal* incest taboo,” implying that the prohibition is less influenced by blood ties and more by sharing the same house.

Similarly, the explanation of the incest taboo provided by the theory of Westermarck, though lacking scientific validity, posits that children who are not biologically related but lived together until the age of six are naturally not sexually attracted to each other in adulthood (Guindi & Read, 2012). It implies that blood ties are not the primary determinant of the taboo, but rather shared upbringing. Furthermore, the fuzziness of the theory is amplified in example (2) where the theory of Westermarck is applied to the situation of sexless couples. The article explains how sexless couples may start feeling like siblings, thus preventing them from feeling sexually attracted to each other again. This illustrates the adaptability of theories surrounding the incest taboo to various contexts. Notably, in example (2), the terms “natural basis” and “it’s natural” contradict the terms “accidentally” and “awkward.”

Example (3) mentions the incest taboo in reference to the prevalence of inbreeding among prehistoric men. The incest taboo may have been implemented to avoid inbreeding that nurtures genetic diseases, but this theory is also scientifically debated (Wolf & Durham, 2005). Thus, there is no consensus on incest and the various reasons why it is taboo. The press reports different interpretations to understand the incest taboo, showing a plurality of theories, emphasising feelings over scientific facts.

Table 5.34 Sample of concordance lines showing different theories on the incest taboo

| N | Np | Date | co-text | node | co-text |
|---|-----|------------|--|----------------------|--|
| 1 | Sun | 06/07/2018 | Yes, what you did was very wrong - but GSA, or genetic sexual attraction , can happen between close relatives like siblings who are not raised together and so not protected by the normal | incest taboo. | |
| 2 | Tim | 18/05/2019 | According to Westermarck, that's the natural basis for the | incest taboo. | If you live with someone for a long period of time and don't have sex together, it's natural you'll begin to feel like siblings. |
| | | | Once the | incest taboo | accidentally kicks in, sex can begin to feel pretty awkward. |
| 3 | Ind | 14/11/2021 | Incest seems to have been | taboo | in early modern human culture, whereas Neanderthal genomes reveal far higher rates of inbreeding which was necessitated by their tendency to live in such isolated communities. |

Finally, there are other mentions of taboo, but on a metalinguistic level. In other words, metalinguistics can refer to linguistic comments on words that are used. Consistent with the earlier findings, abortion is considered the utmost taboo in the British press, and Table 5.35 illustrates how journalists highlight its tabooeness by commenting on their discourse. Phrases

like “*even* discussing abortion is taboo,” “*not* discussed,” “hushed up,” “*rarely* mentioned,” “*keeping* this discussion taboo” demonstrate that linguistic taboos on the topic of abortion are acute, leading to avoidance of the topic. Nevertheless, the examples demonstrate that journalists also attempt to break the taboo by addressing the linguistic challenges of discussing it with metalinguistic comments like “euphemisms prevailed,” “the word + [quotation marks],” and “the subject of.”

Table 5.35 Sample of concordance lines where taboo refers to metalinguistic considerations

| N | Np | Date | co-text | node | co-text |
|---|-----|------------|---|---------------|---|
| 1 | Tel | 19/05/2018 | Here, roads are dotted with graphic anti-abortion signs and residents are far less vocal , as even discussing abortion is | taboo. | |
| 2 | Gu | 15/01/2021 | Sexual pleasure was not discussed , menstruation hushed up . Rape was rarely mentioned . Miscarriage carried a | taboo. | Euphemisms prevailed. |
| 3 | Ind | 27/05/2021 | "By criminalizing women, stigmatizing the subject of abortion and keeping this discussion a | taboo, | we are not saving lives and we are not helping the cause of making abortion obsolete." |
| 4 | Gu | 19/06/2022 | The word "abortion" is no longer so | taboo | (conversations like ours would, she says, have been unimaginable even five years ago). |

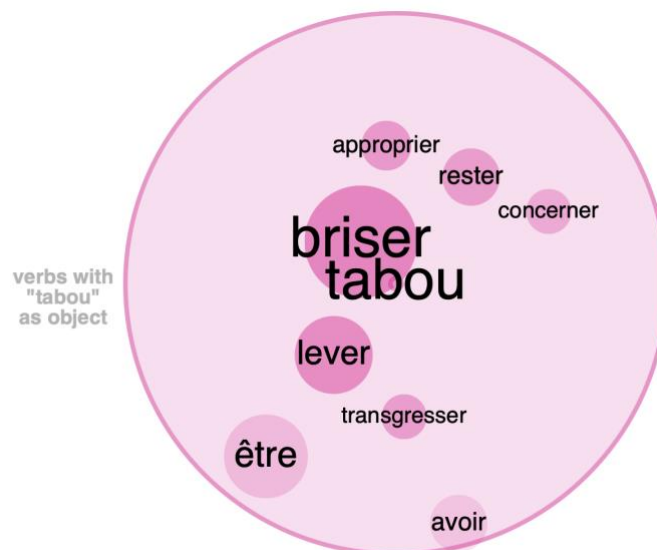
Thus, the metalinguistic comments on abortion highlight its linguistic taboos. Similar linguistic reflections are also present in the French press when discussing the incest taboo. Journalists' comments demonstrate their considerations, difficulties, and caution in addressing such sensitive topics. It may be a journalistic requirement to carefully choose appropriate words and ponder their significance to report facts adequately. The acknowledgement that certain topics, like incest or abortion, are sensitive makes it challenging to address them in discourse. As Jay (2000) notes, the taboo strengthens the taboo. Consequently, linguistic taboos intertwine with social taboos. Sometimes, these topics are preferred not to be discussed at all, despite outward signs in society indicating otherwise. For instance, in example (1), the journalist reports a situation in Texas where inhabitants are more comfortable displaying signs opposing abortion than directly discussing it. Thus, images and signs might convey meaning more easily than words, as the latter may be unspeakable.

3.2. Breaking the taboo by saying the unspeakable

As previously mentioned, the term *taboo* occurs 313 times in the French corpus. The occurrences predominantly refer to the act of breaking the incest taboo. The semantic prosody is neither positive nor negative, as the voices of victims are highlighted for speaking out about their trauma. As illustrated in Figure 5.7, the taboo is embedded in a process of denouncing the prevalence of incest as CSA, with the first two verbs collocating with *taboo*,

being *briser le tabou* (breaking the taboo) and *lever le tabou* (lifting the taboo). Thus, the voices of victims are amplified by acknowledging their bravery, but they are also filtered by the press which comments on their testimonies. Therefore, this section analyses whether the taboo is indeed broken by delving into the mechanisms of incest or whether breaking the taboo remains superficial, limited to the metalinguistic level of language.

Figure 5.7 The most frequent verbs collocating with *tabou* as object in the French corpus



visualization by SKETCH ENGINE

The concordance analysis of the collocation *briser le tabou* (breaking the taboo) reveals that the incest taboo is understood linguistically, as shown in examples (1), (3), and (4) in Table 5.36. In these instances, breaking the taboo means breaking the silence on incest. The conspiracy of silence is crucial in perpetuating incest, as discussed in Chapter 3. Example (2) illustrates the rarer meaning of the incest taboo as a social prohibition, and not as metalinguistic comment. Nevertheless, example (2) demonstrates how the incest perpetrator denies the accusations of incestuous abuse made by several family members and accuses them of “family conspiracy.” The notions of denial and silence are intertwined because speaking out has tangible consequences in families that often implode between those who believe the victims and those who deny intrafamilial abuse (Brey & Drouar, 2022; CIIVISE, 2022, 2023; Dussy, 2021). Such a notion of destroying the family is also mentioned in example (1) with the reported speech of the victim whose “breaking the taboo” has caused “the total explosion of a family.” Interestingly, example (3) provides the testimony of a victim who considers that speaking out is harder than the incestuous abuse, namely that the revelations have a stronger influence than the incest itself. These examples demonstrate that the act of speech, and more specifically the act of denouncing – as noted earlier, *to denounce* is the strongest collocating verb with *incest* as the direct object – is crucial to break the taboo of incest. This is why the notion of metalinguistics sheds light on the silencing mechanisms of incest. Victims voice their abuse by breaking the silence that was imposed on them by the

abuser or the family at large. Breaking the taboo has a cost due to its tangible consequences on families that may implode. As such, language is performative, and is portrayed as powerful and risky.

Table 5.36 Sample of concordance lines of the verb *briser* (break) with *tabou* as object referring to incest

| N | Np | date | co-text | node | co-text | translation |
|---|-----|------------|--|------------------|---|--|
| 1 | O-F | 28/06/2019 | En | brisant le tabou | de l'inceste, l'homme qui est ce mercredi à la barre du tribunal a provoqué « l'explosion totale d'une famille ». | By breaking the incest taboo , the man who is in the dock on Wednesday caused “ the total explosion of a family ”. |
| 2 | Pg | 20/01/2021 | Incapables d'avouer qu'ils ont | brisé le tabou | de l'inceste, ils accusent les victimes - leurs propres petits-enfants, nièces ou neveux - de mentir, ou échafaudent des histoires de complot familial. | Unable to admit that they broke the incest taboo , they accuse the victims – their own grandchildren, nieces, and nephews – of lying , or concoct stories of family conspiracy . |
| 3 | Mt | 13/09/2021 | Elle dénonçait l'inceste dont elle était victime mais à l'époque, | briser le tabou | était plus monstrueux que l'inceste lui-même. | She denounced the incest of which she was a victim but at that time, breaking the taboo was more monstrous than incest itself . |
| 4 | VdN | 17/01/2021 | Au fil des témoignages, courts mais faisant tous état du même traumatisme d'avoir été abusé sexuellement, enfant, par un adulte de la sphère familiale, de nombreux messages de soutien et d'encouragement à | briser le tabou | affluaient. | As the testimonies unfolded, short but all expressing the same trauma of having been sexually abused as a child by an adult of the family circle , many messages of support and encouragement of breaking the taboo poured in. |

Furthermore, the second most frequently occurring verb with *tabou* as the direct object is *lever* in the French corpus, which can be translated by “lift” in the sense of revealing. The sample in Table 5.37 shares similar features with the previous sample shown in Table 5.36, i.e., the mechanisms of breaking or lifting the taboo are described in similar ways. The phrase is used to emphasise the act of speaking out, with the verb “speak” in example (2) or “loosen their tongues” in example (4). Furthermore, there are also descriptions of family implosions as speaking out destabilises the families, with phrases like “tore her family apart” in example

(1) or “eating away at his family” in example (3). The notion of having difficulty to denounce is also present, with the time needed to finally testify emphasised in example (1) as “over thirty years,” or the mention of cutting through, with “she was the first” in example (3). Thus, the French press emphasises the metalinguistics of the incest taboo in exposing the tangible consequences of language regarding the incest taboo: either speaking out and taking the risk of family implosion or remaining silent and keeping the incest taboo unheard of.

Table 5.37 Sample of concordance lines of the verb *lever* (lift) with *tabou* as object referring to incest

| N | Np | date | co-text | node | co-text | translation |
|---|-----|------------|---|------------------------|---|--|
| 1 | Lib | 06/01/2021 | La fille de Bernard Kouchner a mis plus de trente ans à | lever le tabou | sur le drame qui a déchiré sa famille , accusant son beau-père le politologue Olivier Duhamel d'avoir abusé de son frère jumeau . | Bernard Kouchner's daughter took over thirty years to lift the taboo on the tragedy that tore her family apart , by accusing her stepfather , the political scientist Olivier Duhamel, of having abused her twin brother . |
| 2 | O-F | 19/02/2021 | Marielle Hamard parle pour | lever le tabou | de l'inceste. | Marielle Hamard speaks to lift the incest taboo . |
| 3 | Pg | 24/06/2021 | C'est elle qui, la première , | a levé le tabou | de l'inceste qui rongeait cette famille. | She was the first to lift the incest taboo that was eating away at this family . |
| 4 | DL | 22/06/2022 | <i>Cchhhhut !</i> est une pièce grinçante dans laquelle le clown Tomi éclaire des zones sombres de l'être humain et invite à délier les langues pour | lever le tabou | de l'inceste. | <i>Cchhhhut !</i> is a scathing play in which Tomi the clown sheds light on dark areas of the human being and encourages people to loosen their tongues to lift the incest taboo . |

Therefore, the notion of the unspeakable is crucial in understanding incest. Incestuous abuse committed against children can be considered unspeakable or unsayable. In the French corpus, there are 43 occurrences of the term *indicible*, in contrast to 12 occurrences of *unspeakable* and 2 occurrences of *unsayable* in the English corpus. Here, the English language makes a distinction that the French language does not: *indicible* can be translated either as *unsayable* or *unspeakable*. According to the *Oxford Dictionary*,⁶ *unspeakable* defines something that is “not able to be expressed in words,” with a connotation of something “too bad or horrific to express in words.” For the term *unsayable*, the definition states: “not able to be said, especially because considered too controversial or offensive to mention.” Furthermore, Langton (2009) uses the term *unspeakable* to analyse the silence imposed on powerless groups, such as children confronted to sexual abuse. As such, the term *unspeakable*

⁶ Online dictionary, latest update in 2023, version 6.3

was selected over *unsayable* for translating *indicible* because it emphasises the challenge of finding the words on a taboo topic.

Nevertheless, it is worth noting that there might be some “apparent self-contradiction” because saying that something is unspeakable makes it sayable (Ho, 2024, p. 409). However, the examples in Table 5.38 illustrate how the notion of *unspeakable* is used to discuss incest in the French press. The notion is reinforced with additional terms like *horror* in examples (1) and (4), *unconceivable* in example (2), and *unbearable* in example (3). The verbs *hearing*, *conceiving*, and *imagining*, emphasise the difficulty of acknowledging the existence of the taboo. Furthermore, examples (2) and (3) show how discourse on incest euphemises the taboo, with the pronoun *ça* (this) to refer to the daughter’s pregnancy by her dad, or the verb *contourner*, meaning “skirt round” or “circumnavigate” (with words), to comment, encapsulate, and veil the victims’ testimonies in court. Finally, the reported speech of the lawyer in example (4) praises the victim’s bravery in speaking out, despite the costs it represents. Thus, the emphasis on the unspeakable highlights the taboo nature of incest.

Table 5.38 Sample of concordance lines of *indicible* (unsayable) referring to incest

| N | Np | Date | co-text | node | co-text | translation |
|---|-----|------------|--|-------------------------|--|---|
| 1 | L'H | 13/10/2017 | L'autre de | l'indicible, | de l' horreur d'un viol répété jour après jour par un père tout puissant et incestueux , horreur que personne ne sera jamais « formé pour entendre », car entendre l' horreur ne relève pas et ne relèvera jamais d'une « formation ». | The other meaning is the unspeakable , the horror of a rape repeated day after day by an all-powerful and incestuous father , horror that no one will ever be “trained to hear”, because hearing horror is not and will never be a matter of “training”. |
| 2 | S-O | 03/02/2021 | Jusqu'à ce que sa mère conçoive | l'indicible, | l'inconcevable : « Ne me dis pas que c'est ton père qui t'a fait ça . » | Until her mother conceives the unspeakable , the unconceivable : “Don't tell me it's your father who did this to you”. |
| 3 | Mt | 18/12/2021 | « L'insupportable », | « l'indicible », | « l'inconcevable ». Les qualificatifs pour tenter de contourner avec des mots ce que disent avoir vécu les deux enfants de l'accusé se sont succédé pour ce dernier jour de procès. | “ The unbearable ”, “ the unspeakable ”, “ the unconceivable ”. The qualifiers trying to circumnavigate with words what the two children of the defendant say to have experienced |

| | | | | | | |
|---|----|------------|--|--------------------|-------------------|---|
| | | | | | | came one after the other on this last day of trial. |
| 4 | DL | 04/03/2022 | « Imaginez une fillette de 7 ans à qui son père, le sexe en érection devant elle, demande de “sucrer comme une glace” », rappelle l’avocat général qui salue « la force et le courage de cette jeune fille qui a déposé plainte et qui est venue, à la barre d’une audience publique, décrire | l’indicible | horreur ». | “Imagine a 7-year-old little girl to whom her father, with his sex in erection in front of her, asks ‘to lick like an ice cream’”, reminds the lawyer who praises the strength and courage of this little girl who filed a complaint and who came, testifying at a public hearing, to describe the unsayable horror ”. |

Therefore, the French press focuses on the linguistic and social taboos surrounding incest as CSA. Breaking the taboo enables the victims, once they become adults, to reclaim control over a language that has been violated for them. Former victims who become writers elucidate how language is violated for several reasons (Angot, 2013; Sinno, 2023). Firstly, the name of a parent, like “daddy,” becomes a non-sense due to the kinship transgression, rendering the most basic words meaningless. Secondly, victims are denied the right to speak out, leading to the negation of incestuous abuse. When incest is silenced, its existence is denied, as evidenced by the frequency of the term *silence*, appearing 673 times in the French corpus. Similar to the term *tabou*, the first collocating verb is *briser* (break). As such, it indicates that *silence* is synonymous with *taboo* in the sense that breaking the taboo is breaking the silence on incest. It draws us back to the observations on the definitions of incest in Chapter 2. When incest is defined as a marital prohibition, it is perceived as a social taboo (Lévi-Strauss, 1949). In other words, incest is a universal rule proscribing sexual relationships between family members. However, when incest is considered intrafamilial sexual abuse, involving underage children, it is primarily understood as a linguistic taboo. In other words, the victims are coerced into silence and do not disclose the transgression of the universal taboo. When victims finally find their voices, they challenge language for naming the truth. The emotional cost of disclosing familial abuse demonstrates that language is at stake of power, as such revelations can have devastating consequences. Language, being performative, can have tangible repercussions. However, it remains relevant to question whether this focus on language has always been prevalent in the French press or if it was prompted by the high-profile scandal happening in 2021, which brought the incest taboo to the forefront. This question is addressed in the following section.

3.3. Newsworthiness of the incest taboo in 2021: the diachronic analysis of the French corpus

In 2021, Camille Kouchner, the daughter of Bernard Kouchner, former French Minister and co-founder of Doctors without Borders and Doctors of the World, sparked a public scandal by publishing *La Familia grande* on January 7 (Kouchner, 2021). The book breaks a longstanding silence in her family, revealing the incestuous abuse her twin brother endured in his teenagehood at the hands of their stepfather, Olivier Duhamel. The book, in the form of a testimony, delves into the mechanisms of denial within the family, with her mother refusing to acknowledge the abuse and her brother remaining silent. Kouchner explains how the family, relatives, and society at large enable incestuous abuse to persist for years. Duhamel is never named officially in the book, but he is recognised in the narrative and was later identified by the French press. He was a political scientist teaching at the very prestigious Parisian university Science-Po, renowned for its political science curriculum.

Several factors converge to render this story a newsworthy scandal in the French press, due to three news values (Bednarek & Caple, 2019; Harcup & O'Neill, 2001, see Chapter 4). The first news value is celebrity. The protagonists belong to what could be termed the Parisian establishment. Olivier Duhamel was an influential man in politics, and the name Kouchner is well-known in France. The second news value is unambiguity, as the revelations are documented in a book, providing insights into the mechanisms of incest. The third news value is continuity. Following the #MeToo movement in 2017 when thousands of women spoke out against sexual harassment in the workplace, French society became increasingly aware of endemic sexual violence, empowering victims to voice their experience. Thus, the disclosure of incestuous abuse within a prominent family continued the ongoing discussion on sexual violence in the French press. Another indication of this continuity was observed on the social media platform then called Twitter, where the hashtag #MeTooInceste gathered 80,000 messages from victims recounting their abuse in two days (Carrive, 2021; de Foucher, 2021).

Therefore, we can wonder whether this scandal has changed the newsworthiness of the incest taboo for the French press. In other words, has the press focused more on the incest taboo as CSA due to this scandal, or was this focus always predominant? Are there diachronic absences in the French corpus?⁷ As outlined in Chapter 4, 40 % of the French corpus consists of articles published in 2021, representing almost half of the corpus. Consequently, the French corpus was further divided to compare before and after 2021. The first subcorpus, named Corpus 17-20, comprises articles published between 15 October 2017 to 31 December 2020. The second corpus, named Corpus 21-22, comprises articles published between 1 January 2021 to 14 October 2022. Corpus 17-20 contains 1,309 articles, amounting to 655,719 words, and Corpus

⁷ This diachronic analysis is only conducted on the French data, as the British press did not cover this high-profile scandal or any other type of scandal (see Chapter 4, Section 3, for the yearly and monthly proportional numbers of newspaper articles of both corpora).

21-22 contains 2,037 articles, amounting to 1,184,082 words. This indicates that for a shorter time span (22 months), Corpus 21-22 is twice the size of Corpus 17-20 (38 months).

This initial observation suggests that the publication of the book had a significant impact on the French coverage of the incest taboo, making it newsworthy. Thus, the aim is to investigate whether the news was covered differently before or after 2021. For this purpose, Corpus 17-20 is compared to Corpus 21-22, and *vice versa*, to determine the presence or absence of certain words before and after 2021. As such, positive keywords, present in the corpus, and negative keywords, absent in the corpus, indicate what differs from one corpus to another. A positive keyword is defined as “a word that occurs significantly more often in one corpus, compared to the other” in contrast with a negative keyword that “occurs significantly less often in one corpus, compared to the other” (Baker, 2023, p. 167). Table 5.39 demonstrates that after 2021, the primary focus of the French press is the incest scandal involving Olivier Duhamel, since the positive keywords of Corpus 21-22 refer to it, whereas those keywords are absent in Corpus 17-20.

Table 5.39 Positive keywords in Corpus 21-22 compared to Corpus 17-20

| N | Keywords | Translation |
|-----------|-----------------|---|
| 1 | CIIVISE | Independent Commission of Inquiry on Incest and Sexual Violence committed to children |
| 2 | familia | familia |
| 3 | Mion | Mion, ex-Director of Science Po |
| 4 | Duhamel | Duhamel |
| 5 | Dupond-Moretti | Dupond-Moretti, French Minister of Justice |
| 6 | MeTooInceste | MeTooInceste |
| 7 | Kouchner | Kouchner |
| 8 | sciences | sciences |
| 9 | FNSP | National Foundation of Political Science |
| 10 | Gorce | Gorce |

As can be observed, the fifteen keywords pertain to the family names involved in the scandal, such as Duhamel and Kouchner, as well as names associated with the political ramifications of the revelations, such as Mion, Dupond-Moretti, and FNSP. It is worth noting that the decision to establish a French Commission of Inquiry against Incest and Sexual Violence (CIIVISE) with a two-year mandate preceded the scandal. However, the scandal rendered the commission more relevant, leading to increased coverage by the press. Therefore, the presence of these keywords in Corpus 21-22, as opposed to their absence in Corpus 17-20, indicates the significant impact of the scandal, which the press extensively covered in 2021. Conversely, the keywords present in Corpus 17-20, but not in Corpus 21-22, relate to plays featuring historical figures where incest is staged, like *Borgia* (Borgia), *Lucreèce* (Lucrezia) or *Sénèque* (Seneca), often accompanied by the months of the performances (October, January). Thus, the scandal appears to have contributed to increased awareness of incest as CSA. However, comparing positive and negative keywords, in other words, present and absent keywords from one

corpus to another, necessarily highlights differences. This is why the frequency lists are observed to contradict the topicality of each subcorpus, as seen in Table 5.40.

Table 5.40 The fifteen most frequent lexical nouns with their translation in each subcorpus

| N | Corpus 17-20 | Translation | Corpus 21-22 | Translation |
|----|--------------|------------------|----------------|---------------|
| 1 | an | year | an | year |
| 2 | h | hour (a.m./p.m.) | victime | victim |
| 3 | enfant | child | enfant | child |
| 4 | victime | victim | inceste | incest |
| 5 | deux | two | violence | violence |
| 6 | femme | woman/wife | femme | woman/wife |
| 7 | fille | daughter/girl | deux | two |
| 8 | père | father | fait | fact |
| 9 | homme | man | fille | daughter/girl |
| 10 | fait | fact | viol | rape |
| 11 | viol | rape | père | father |
| 12 | cour | court | homme | man |
| 13 | inceste | incest | famille | family |
| 14 | mère | mother | année | year |
| 15 | violence | violence | mère | mother |

Interestingly, the subcorpora share the same top fifteen words, except for *hour* and *court* in Corpus 17-20 and *family* and *year* in Corpus 21-22. Another slight difference is the frequency of the term *incest*, which occurs more often in Corpus 21-22, with 3,003 occurrences compared to 600 occurrences in Corpus 17-20. This means that incest occurs nearly three times more frequently in Corpus 21-22 when proportions are considered, indicating that the term is used more often in the articles. Nevertheless, the lexical nouns are largely consistent across both corpora, indicating that the French press has covered incest as CSA before and after 2021. Despite the fact that 40% of the French corpus consists of articles published in 2021, indicating the scandal's impact, the portrayal of incest by the press remains consistent. In other words, the newsworthiness of the incest taboo increased in 2021, but the press coverage consistently frames incest as CSA.

However, there is a clear emphasis on the notion of speaking out after 2021. It has been previously noticed that silence and denial are crucial elements in understanding incest. The newsworthiness of the incest taboo focuses on breaking the silence. Indeed, articles in the corpus highlight how the voices of the victims are finally being heard. The newsworthiness of breaking the incest taboo is expressed through the mention of “libération de la parole” (liberation of speech). Notably, the term *parole* (speech/word) occurs 1,001 times in Corpus 21-22, and 343 times in Corpus 17-20, which means that the term is six times more frequent in the second subcorpus than the first, when proportions are considered. This suggests that the discourse on incest in the French press has consistently portrayed incest as CSA, but with an emphasis on silence and speech in 2021.

As such, the concordance analysis demonstrates that the phrases composed with *parole* (speech) are similar in both subcorpora. The press covers “groupes de paroles” (speech group) where people can share their experience of sexual abuse. Furthermore, the possessive pronouns qualifying *parole* are *sa* (his/her), *leur* (their), and *votre* (your). Thus, the act of speaking is embodied by individuals. However, it is worth noting that the deictic *cette* (this), in *cette parole* (this speech), only occurs in Corpus 21-22. The phrase *cette parole* indicates a change of paradigm in the way the victims are listened to in 2021. As mentioned above, the emphasis on speaking about incest is more acute after 2021. Likewise, “this speech” suggests that speaking about incest is not just a personal or individual matter but rather a social phenomenon involving awareness towards the taboo of incest. Thus, *cette parole* personifies the victims’ voices, focusing on their social effects. It is a more metalinguistic perception of the speech as it enables political and social changes, as Table 5.41 illustrates.

Table 5.41 Sample of the concordances *cette parole* (this speech)

| N | Np | Date | co-text | node | co-text | translation |
|---|-----|------------|--|----------------------|---|--|
| 1 | Pg | 25/01/2021 | | Cette parole | libérée a conduit Emmanuel Macron à poster samedi un message vidéo de soutien aux victimes . | This freed speech led Emmanuel Macron to post a video message of support to the victims on Saturday. |
| 2 | L'H | 18/02/2021 | Relayée par les réseaux sociaux , formidable caisse de résonance, | cette parole | s'affranchit de toutes les barrières qui jusqu'à maintenant l'empêchaient de s'exprimer . | Spread by social media , a great echo chamber, this speech breaks free from all the barriers that were preventing it from speaking out until now. |
| 3 | DL | 23/02/2021 | Maintenant, l'institution est dans une phase où elle doit décider ce qu'elle fait de | cette parole. | | Now, the institution is in a phase where it must decide what to do with this speech . |
| 4 | Lib | 26/02/2021 | Il s'agira à la fois d'offrir un accompagnement de l'ordre du soin et de partir de | cette parole | pour construire une meilleure protection des enfants . | It will involve both providing support in terms of care and using this speech to build better protection for children . |
| 5 | O-F | 23/08/2021 | Ce film a créé entre nous deux une sorte de sororité dans laquelle nous partageons et nous portons | cette parole | interdite, ce silence qui crie . | This film created some kind of sorority between the two of us where we share and carry this forbidden speech, this silence that screams . |

| | | | | | | |
|---|-----|----------------|--|---------------------|------------------------|---|
| 6 | Fig | 20/09/ 2021 | Elle témoigne de la prise en considération de | cette parole | par la société. | It proves the consideration of this speech by society. |
|---|-----|----------------|--|---------------------|------------------------|---|

4. Summary

This, this chapter has highlighted the different perspectives chosen by the French and British press on the incest taboo. The press coverage of this taboo, defined as universal, seems to be culturally determined. The British press primarily defines the incest taboo as a social taboo, since the depicted relationships are consensual but transgress a social prohibition. Conversely, the French press primarily defines the incest taboo as a linguistic taboo, shedding light on the prevalence of silenced intrafamilial CSA. As such, the French press highlights the victims' voices who break the silence on the incest taboo through metalinguistic comments. Remarkably, such depictions in the press align with the legal definitions of incest in each country. British law prohibits consensual relationships, and this type of relationship is predominantly portrayed in the press. Conversely, French law defines incest as rape and sexual assault, and incest is predominantly portrayed as CSA. All in all, it is worth considering how deeply entrenched cultural perceptions shape media representations. Thus, it calls into question the unfathomable nature of cultural differences.

Furthermore, the relevance of the traditional dichotomy between broadsheets and tabloids in the British press landscape can be questioned regarding a sensitive topic like incest. The British press reports incest when it is sensational, and similar linguistic patterns are found in both types of newspapers. Fictional incest is a predominant focus across newspapers where the eroticism of the incest taboo influences the representations of incest and marginalises its social implications. The sensationalism of news items is characteristic of the tabloidisation of the press (Brown et al., 2018). The broadsheets report incest stories when they are sensational, involving murders or celebrities. For this reason, incest is not discussed for itself, but is presented as a plot element in a list of other mischiefs. Furthermore, this fictional representation of incest is created by authors and directors, but not narrated by former victims. A broader question that could be addressed from this observation is whether generally, there are fewer first-person narratives on incest in the British literary landscape, or whether the press chooses to focus only on the fictive and televised stories. Another research could inquire how incest is culturally portrayed in fiction, both in France and in the UK, and how it determines its newsworthiness.

Finally, this chapter has reflected on the notion of discursive absence. The absent coverage of the incest taboo as intrafamilial CSA in the British press gave rise to different approaches to find this known absence. However, this investigation was challenged by the limits of finding something that is textually absent, because socially tabooed. The method of comparing two corpora to analyse the press coverage of a sensitive topic proves to be a reliable means of analysing discursive absence. Notably, the English corpus comprises a significant proportion of articles related to the abortion ban in the US with no exception for rape or incest. This

suggests that if not for the abortion ban, there would be considerably fewer articles featuring the lemma *incest**, underscoring the strong taboo surrounding the topic in the British press. It also demonstrates the absence of the topic, indicating the utmost form of taboo, that is silence. For this reason, the next chapter delves further into the linguistic avoidance strategies of the incest taboo.

Chapter 6: Searching for Absence

And how can we analytically engage with something that may not be linguistically present and argue the existence of the non-existent voice?

(Strand, 2018, p.126)

The previous chapter contrasted the findings of the French and English corpora built with the lemma *incest**. The results suggested that the perceptions of incest differ between the UK and France, and that their respective legal definitions of incest are mirrored in the media discourse. While the French press and law primarily define incest as intrafamilial CSA, the British press and law primarily portray incest as transgressive yet consensual relationships. As such, the absence of depicting incest as intrafamilial sexual abuse in the British press raised methodological challenges. It was called into question whether the corpus linguistic tools reached their limitations for hidden occurrences of incest as CSA, or whether this depiction is simply absent in the British press.

Therefore, this chapter aims to further investigate taboo language regarding incest.⁸ As such, it looks for occurrences of incest in the press that are covered and discussed indirectly. In doing so, another search query was generated on the assumption that *incest* is an avoided term due to its taboo nature. For this purpose, this chapter delves into the creation of the search query “abuse AND father” that required the development of an iterative approach. Furthermore, this search query raises cross-linguistic challenges regarding the semantic prosodies of *abuse*: *abus* between the English and French languages. Consequently, the filtering process of the search query is described at length to explain the corpus linguistic approach to search for the unspeakable. Finally, the notions of ambiguity and vagueness related to taboo language are discussed further.

1. Developing an iterative approach

1.1. The search query built on known absence

As observed in the previous chapter, incest as CSA is scarcely reported in the British press, whereas the prevalence of incestuous abuse is more common than acknowledged (Brownmiller, 1993; Dussy, 2021; Justice & Justice, 1980, see Chapters 2 and 5). The hypothesis based on this observation is that newspapers might cover incestuous stories without clearly

⁸This chapter expands on the article “Searching for the unspeakable: An iterative approach to designing a corpus of texts about a taboo topic” written with Gavin Brookes and published in *Research Method in Applied Linguistics* (2024), and the article “Taboo language and signalling nouns: When newspapers substitute, rename, or define taboo topics” written with Andrzej Łyda and published in *Orbis Linguarum* (2024).

stating that it is incest. In other words, an article might report intrafamilial sexual abuse without qualifying it as incestuous. The underlying sociolinguistic assumption is that *incest* is a taboo word as it is not dared to be openly discussed. Drawing on this hypothesis, this chapter seeks to determine whether a corpus can be built on a taboo topic where the keyword is absent.

Thus, the aim is to investigate whether data can be collected from an absent keyword in corpus linguistics. As previously mentioned, corpus linguistics is a methodology that enables the collection of a large amount of data and reduces researchers' bias by using computational tools (Baker, 2023; Tranchese, 2023, see Chapter 4). One way to collect the data is to build a corpus through a search query. Selecting the search terms to build a corpus is critical for determining which texts are included or excluded (Baker, 2023; Baker et al., 2013). The choice of search terms is especially relevant in the context of media discourse representation, as it influences the retrieval of newspaper articles by software (Baker & Levon, 2015; Brookes & Baker, 2021; Freake, 2012; Gabrielatos, 2007; Taylor, 2014). Thus, conducting corpus linguistic research involves considering the delimitation of a topic in what may be present in a text.

However, in this chapter, the selection criterion to build the corpus is absence. Therefore, building a corpus on absence raises methodological challenges as one of the limitations of corpus linguistics is to be "primed to look at what is *there*" (Gillings et al., 2023, p. 45, emphasis in original). As discussed in Chapter 4, tools like collocations and concordances are designed to find linguistic patterns in a corpus and do not seem suitable for finding absence (Partington, 2014). As such, investigating absence can be considered one of the dusty corners of corpus linguistics (Taylor & Marchi, 2018, see Chapter 4). The first challenge lies in the paradoxical nature of absence "because, by definition, it cannot really be said to be anywhere" yet "it is everywhere" (Duguid & Partington, 2018, p. 39). The second challenge is the risk of overinterpreting because "[h]ow do we come to notice absences?" (Schröter & Taylor, 2018, p. 1). There might be missing elements in a text indicating censored or taboo topics, but the risk lies in overinterpreting what is absent in a text, and in projecting the researchers' expectations into it.

Furthermore, a distinction was made between absence and silence in Chapter 3 as the incest taboo is characterised by its silencing. In this thesis, silence is primarily perceived as the coercive process of not saying. However, silence should not be perceived only negatively, since silence is meaningful, and sometimes even respectful (Jaworski, 1993). The question of respectful silence can be raised in the case of incest due to its traumatic nature. In contrast, absence is here defined as the consequence of silence: a topic is silenced due to its taboo nature; thus, it is absent in discourse. The discursive absence actively sought by the researcher can be termed the known absence, i.e., "linguistic feature you are searching for and simply want to know whether or not it is in the corpus" (Partington, 2018, p. 39). In other words, *known absence* depends on the linguistic feature or political issue that the researcher is aware of and is searching for. In Chapter 4, Section 2, methodologies addressing the challenges of finding absence were discussed, and Chapter 5 applied the methodology of comparing

corpora to see if some elements are missing from one dataset in comparison with another. This chapter aims to provide an additional methodological solution: Developing an iterative approach to generate a relevant search query looking for the unspeakable.

A relevant search query must be initially determined to select the targeted articles, i.e., articles covering incestuous stories without mentioning incest. For this purpose, alternative terms must be determined. Thus, the first step of the iterative approach is to test out the term *abuse*. The analysis conducted in Chapter 5 revealed that *abuse: abus* is often used in descriptions of incest, both in the French and British press. The hypothesis is that *abuse* is a general term, broad enough to encompass several meanings and retrieve numerous articles. However, *abuse* may also be too vague. Therefore, a second search term is added to refine the search query, *father: père*. This additional term, *father: père*, is a critical choice. Statistics show that 95% of incest perpetrators are men and 78% of victims are girls. The most common instances of incest are committed by fathers or step-fathers (Gilgun, 1995, p. 268; Yildirim et al., 2014). The lemma *father* includes terms like *step-father* and *grand-father* and the lemma *père* includes the equivalent terms *beau-père* et *grand-père*. However, it should be noted that choosing the search term *father: père* may strengthen stereotypes regarding incest. Mothers can also be incest perpetrators, but victims feel embarrassed to have been assaulted by a woman and tend to change the aggressor's gender when they reveal the sexual abuse (Hetherington, 1999). Assaults committed by women are underestimated in statistics, and are in fact an even greater taboo (Elliott, 1993). Hence, it is more likely to find articles related to father than other family members. For these reasons, the search term *father: père* was selected along with the term *abuse: abus*. Thus, the search queries “abuse AND father” and “abus AND père” narrow the resulting corpora but also what they may represent as they focus on father figures. However, this narrowing of focus is deemed necessary to build relevant corpora.

In doing so, the corpora are collected with the search terms “abuse AND father” in the British press and “abus AND père” in the French press on the news archive Lexis Nexis (see Chapter 3, Section 3, for further information on Lexis Nexis). The English interface of Lexis Nexis works with the Boolean search that uses AND, OR and NOT to define the search. Hence, in the search queries “abuse AND father,” the Boolean search indicates to the database that both *abuse* **and** *father* must be present in the retrieved articles. It is also interesting to note that the English interface of Lexis Nexis affects the French search query. Consequently, “AND” cannot be translated into the French conjunction “ET” since the software is coded in English. This is why the French search query is written as “abus AND père.” Lexis Nexis was used in its default settings. The time period was the same as the one for the corpora built with the lemma *incest**, i.e., from 15 October 2017 to 14 October 2022, covering the emergence of the #MeToo movement up to the first data collection day. In the British press, articles are scraped from the same four broadsheets – *The Daily Telegraph*, *The Guardian*, *The Independent*, *The Times*, and four tabloids – *Daily Mail*, *Daily Mirror*, *The Express*, *The Sun*. In the French press, articles are retrieved from the same five national newspapers – *La Croix*, *Les Échos*, *Le Figaro*, *L'Humanité*, *Libération*, and eight regional newspapers – *Le Dauphiné libéré*, *La Dépêche du*

Midi, L'Est républicain, La Montagne, Ouest-France, Le Progrès, Sud-Ouest, La Voix du Nord. At this first stage, the English search terms “abuse AND father” produce 23,015 article results and 28,187,466 words, while the French search terms “abus AND père” produce 2,796 results and 2,029,732 words. As observed, even though this first search query is similar in French and English languages, the English corpus retrieved eight times more articles than the French corpus and close to fourteen times more words than the French corpus. The notion of semantic prosody is explored in greater detail in the following section to shed light on this discrepancy.

1.2. The cross-linguistic challenges of semantic prosody

As discussed in Chapters 4 and 5, semantic prosody “is an important consideration in identifying functional equivalence” (Taylor, 2014, p. 374) since similar words between two languages, like *abuse: abus*, may refer to different realities. For this reason, “[c]ross-linguistic comparisons, then, require in-depth analysis of translation equivalents and their semantic prosodies in order to make the analysis balanced” (Vessey, 2013, p. 14). The terms *abuse: abus* were purposefully chosen for their vagueness, which can encompass different meanings, thus not limiting the retrieval of articles. However, this vagueness seems stronger in the English language than in the French language, since the number of collected articles through this search query is eight times higher in the English corpus. In order to determine the semantic prosody of *abuse: abus*, their collocates are examined. For this purpose, both corpora are uploaded on the software Sketch Engine (for further explanations on Sketch Engine, see Chapters 4 and 5).

Thus, the collocates of *abuse* in the English corpus are examined to determine the different meanings, and whether these relate to the topic of incest. As previously noted, collocates allow the researcher “to examine how words are used in context (as opposed to dictionary definitions), and also which contexts they do *not* occur in” (Baker, 2006, emphasis in original). To observe how *abuse* is categorised in the data, the collocates are sorted on the left side of *abuse*, since these premodifiers function as adjectives. As such, the frequent collocates can be broken down into ten categories. The prefix **anti** constitutes a category relating to campaigns fighting abuse: *anti-child, anti-clerical, anti-gay, anti-immigrant, anti-Irish, anti-online*, and *antisemitic* abuse. The other categories refer to **religion** – *church, clergy, clerical, dowry, honour-based* abuse; **drugs** – *alcohol, cannabis, drug, cocaine, opioid, substance* abuse; **children** – *child, child sex* abuse; **economy** – *economic, financial* abuse; **elder people** – *elder, elderly* abuse; **physical** abuse; **racism** – *racial*, racist* abuse; – and **verbal** abuse, with colligations such as *hurl* and *yell*. Thus, this concordance analysis confirms the hypothesis that the search term *abuse* is vague enough to encompass varied meanings. Nevertheless, the most frequent collocations, and deemed particularly relevant to the research question, relate to **sex*** (*sex, sexual*) abuse.

In order to contrast these findings and understand the discrepancy in the number of articles retrieved between the French and English corpora, the same collocate analysis is conducted

on the French corpus. However, due to the different grammatical construction, the collocates of *abus* are examined on its right side, as the adjectives are in a post-modifier position in French. Interestingly, only two predominant semantic fields are distinctive. The first one relates to **sex*** *abuse*. The brief observation of the concordance lines suggests that *abus* is mainly used in reference to sexual violence in the Catholic Church, where *père* (father) refers to priests. The second prevalent semantic field pertains to the **law**, with terms like *abus de faiblesse* (abuse of weakness), *abus de confiance* (breach [lit. abuse] of trust), *abus de pouvoir* (abuse of power), *abus d'autorité* (abuse of authority), *abus de biens sociaux* (misuse [lit. abuse] of company assets), *abus de droit* (abuse of process). The only exception from these two predominant semantic fields is the recurring collocate *abus d'alcool* (alcohol abuse), also found in the English corpus but categorised as drug abuse.

Nevertheless, the predominant collocation in the French corpus is *abus sexuel*, suggesting that the semantic prosody of *abus* is more strongly associated with sexual violence in French than in English. For instance, terms like *elderly abuse* or *racist abuse* are absent in the French corpus because equivalent collocations do not exist in the French language. The collocation *racist abuse* can be translated by *insultes racistes* (racist insults) or by *racisme* (racism). Similarly, *elderly abuse* can be translated by *âgisme* (agism). Nonetheless, the term *abuse* in English encompasses various meanings that do not exist in French. Therefore, building comparable corpora in two different languages raises methodological challenges from the beginning, as the semantic prosodies of a term that appears identical in both languages refer to different realities in each. Thus, the search query must be adapted to each language to build comparable corpora. This is because the use of the same search queries could produce incomparable results. Therefore, the next step is to identify their respective topicality to determine how the search query could be refined.

For this reason, the next step in the iterative approach is to filter the search query. Thus, the topicality of the corpora should be determined to understand which articles could be selected or discarded to refine the search query. Although a brief perusal of the French corpus suggests that most French articles cover sexual abuse within the Catholic Church, a closer investigation is needed. For this purpose, the frequency list of the corpus built through the search query “abus AND père” is analysed. As in Chapter 4, the frequency of lexical nouns is examined to determine the topicality of the corpus. The ten most frequently occurring lexical nouns confirm the preliminary observation: the corpus focuses on sexual abuse committed within the Catholic Church, with *père* (father) referring to priests, as seen in Table 6.1.

Table 6.1 The top ten lexical nouns in the French corpus “abus AND père”

| N | Noun | Translation | Raw frequency | Frequency per million |
|---|--------|-------------|---------------|-----------------------|
| 1 | an | year | 7,088 | 2,890.87 |
| 2 | père | father | 5,129 | 2,091.89 |
| 3 | église | church | 4,485 | 1,829.23 |
| 4 | abus | abuse | 4,327 | 1,764.79 |

| | | | | |
|----|---------|-----------|-------|----------|
| 5 | victime | victim | 4,291 | 1,750.10 |
| 6 | droit | right/law | 3,939 | 1,606.54 |
| 7 | deux | two | 3,535 | 1,441.77 |
| 8 | enfant | child | 3,251 | 1,325.94 |
| 9 | prêtre | priest | 3,050 | 1,243.96 |
| 10 | homme | man | 2,813 | 1,147.30 |

Further down the list, other terms pertaining to sexual abuse within the Catholic Church are recurrent: *évêque* (bishop) is 17th in the list, with 2,038 occurrences, *diocèse* (diocese) is 26th, with 1,485 occurrences, *François* (Francis) and *Pape* (Pope) are respectively 29th and 30th, with 1,409 and 1,398 occurrences. Therefore, due to their relatively high frequencies, they are added to the search query as excluded terms. Although the denunciation of CSA occurring within the Catholic Church is on the same continuum of sexual violence against children, the prime focus of this research is to analyse its occurrence within the families. Thus, the terms *église* (church), *prêtre* (priest), *bishop* (*évêque*), *diocèse* (diocese), *Pape* (Pope) and *François* (Francis), are introduced in the search query with the conjunctions “AND NOT” to indicate to Lexis Nexis **not** to retrieve articles comprising these terms, even when they include *abuse* and *father*. Therefore, the search query should discard the unwanted articles covering CSA within the Catholic Church but retrieve articles mentioning sexual abuse in a non-religious context. As a result, the second search query, shown in Table 6.2, produced 1,387 results, amounting to 942,092 words. Hence, the second search query divided by two the number of articles and the number of words compared to the first search query. However, to test the efficiency of this search query, 100 articles are closely read, as explained in the following section.

Table 6.2 The second search query to filter the results in the French corpus

| |
|---|
| père AND abus AND NOT église AND NOT prêtre AND NOT évêque AND NOT diocèse AND NOT pape AND NOT françois |
| father AND abuse AND NOT church AND NOT priest AND NOT bishop AND NOT diocese AND NOT pope AND NOT françois |

Similarly, the search query in English needed to be refined. The frequency list of the English corpus is observed to determine other candidate lexical words that could be used as search terms. The ten most frequent lexical nouns are *child*, *year*, *abuse*, *father*, *family*, *time*, *people*, *woman*, *life*, *court*. Interestingly, the highest frequency term is *child* and it seems relevant to potential reported stories on incest, as incest may refer to intrafamilial CSA between an adult and an abused underage child. Therefore, *child* is considered an interesting additional search term. However, due to the diverse semantic prosody of the term *abuse*, three search queries are tested before deciding how to refine the search query. Based on the previous findings, the following search queries are generated to identify the most relevant one:

- (a) child W/5 abuse AND father (2,660 articles)
- (b) abuse AND father AND child (14,684 articles)
- (c) abuse AND father AND sex* (10,041 articles)

The search query (a) “child W/5 abuse AND father” includes texts related to *child abuse*, *child sex/sexual abuse*, or *child sexually abused*, as well as other linguistic patterns where *child* and *abuse* are close collocates with a separation of 0 up to 5 words. The search query (b) “abuse AND father AND child” implies that the terms can appear in any position in the article, contrary to (a). The final search query, (c) “abuse AND father AND sex*” should restrain the meaning of abuse with the lemma *sex**. The wildcard (*) on Lexis Nexis includes all the forms derived from *sex*, thus encompassing *sex*, *sexual*, and *sexually*. In order to determine the most efficient search query, 100 articles of each corpus are closely read, as described in the following section.

2. Optimising the search query

The next step of this iterative approach is to refine the search query, keeping in mind the specificity of each corpus. As observed, despite the seeming synonymy of the search queries “abuse AND father” and “abus AND père” for building the English and French corpora, the amount of data differs significantly from one language to another. Therefore, the new search queries were tested out by closely reading 100 articles of each sample. This step consisted of determining the relevance of the search queries. For an article to be deemed relevant, incest should be referred to either explicitly or implicitly. Terms such as *touch**, *molest**, *rape*, or *sex* abuse** are considered characteristic of incestuous abuse. Thus, the articles were broken down into three categories: a YES category – articles are relevant, a NO category – articles are not relevant, and a MAYBE category – articles are ambiguous, i.e., difficult to categorise into the two other categories. The definition of incest determines if an article belongs to the YES, NO, or MAYBE categories. In this thesis, a relationship is considered incestuous whenever it occurs between family members. The definition of family includes biological, adoptive, foster, and alliance relatives (see Chapter 2).

2.1. The filtering process in the English corpus

As mentioned above, three potential search queries are considered to optimize the search in the English corpus. To determine the most relevant search query, 100 articles are closely read in each subcorpus, (a), (b), and (c):

- (a) child W/5 abuse AND father (2,660 articles)
- (b) abuse AND father AND child (14,684 articles)
- (c) abuse AND father AND sex* (10,041 articles)

Articles assigned to the YES category included those which contained an explicit mention of sexual abuse by a relative. In examples (1) and (2) below, it can be observed that the word *incest* is not used (as is also the case for the rest of these articles), but incestuous abuse is clearly referenced through the perpetrator of the abuse being identified as a family member. Meanwhile, the sexual nature of the abuse is explicitly specified. Interestingly, example (1) shows that although incestuous abuse was committed by the uncle, the article was retrieved

because the term *father* appears subsequently. This demonstrates that the search query can collect articles mentioning other family members perpetrating incestuous abuse.

- (1) Dyer revealed the **sexual abuse** inflicted on him as a **child** by an uncle, in whose care he had been left. [...] She and his **father**, Leroy Dyer, had separated.

Daily Mail, 12 February 2018

- (2) Connolly denied he was burdened with memories of his **father's sexual abuse** of him as a **child**.

Express, 30 November 2018

On the other hand, the terms *abuse* and *father* could occur within the same article but be used in different contexts. In such cases, there was no identifiable mention nor implication of incestuous acts in the reporting. Example (3) demonstrates media coverage, following the #MeToo movement, denouncing sexual abuse in Hollywood, where *father* is used to emphasise a large age gap between the perpetrator and victim. In other cases, *father* could refer to fathers worried about their children, such as their online activity increasing their risk of being exposed to sexual abuse, as in example (4). In this case, *online sexual abuse* refers to paedophilia (this phrase is a frequent substitute for paedophilia, see Chapters 3 and 5). Finally, example (5) demonstrates how the phrase *child abuse* could be constructed as equivalent to forcing children to wear protective face masks, as a way of negatively appraising COVID-19 prevention measures. Thus, the terms *abuse* and *father*, *child* and/or *sex**, can occur together, but refer to completely different situations. Such cases were allocated to the NO category.

- (3) Sir Tom Jones has said the music industry suffers the same issues of **sexual abuse** and harassment that are alleged to have taken place in Hollywood. [...] Headey claimed she tried to laugh it off, telling Weinstein he was old enough to be her **father**.

Independent, 19 October 2017

- (4) Speaking as he met tech giant bosses in the US to get them to do more to tackle **online child abuse**. [...] **The father-of-four** revealed that to safeguard his three children who were still under 18, "I know all of the apps they have on their phone."

Daily Telegraph, 9 November 2018

- (5) "I think his **father** exploited him [...]." The host was Tucker Carlson, whose smirking sympathy for the white supremacist "great replacement" theory, and insistence that making kids wear face masks is "**child abuse**", have made him the true heir to Donald Trump as the rightwing conspiracy-theory king.

Guardian, 30 April 2021

Finally, some articles were considered too ambiguous to be classified either in the YES or NO categories. This ambiguity arose due to vagueness regarding aspects of the story being reported. For this reason, the category MAYBE was created to consider these complex and illustrative patterns of linguistic avoidance strategies. This vagueness could result from the

type of abuse being reported, i.e., the abuse is not clearly said to be sexual in nature, as examples (6) and (7) illustrate. Interestingly, in example (7), *abuse* collocates with *domestic* and *drug*, but the mention of *non-molestation order* is not explicitly explained.

- (6) His autobiography, *Windswept and Interesting*, is out this week, though writing it had been difficult, Connolly said, particularly when he addresses the **abuse** he suffered as a child from his **father** and his **aunts**.

Times, 13 October 2021

- (7) Inspectors reviewed the records for one young **child** who was not linked to her **father** in the system, despite his history of **domestic abuse** and **drug use** and the fact he is subject to a **non-molestation order**.

Independent, 21 February 2022

Vagueness could also occur when sexual abuse is said to have occurred, but there is no direct link between the victim and the mentioned family member, which makes it difficult to determine if it is incest, as in examples (8) and (9):

- (8) Speculating as to what drove him to his actions, the film-maker revealed he had been **sexually abused** as a **child** and a teenager, and said **his father** had left his mother when he was young.

Independent, 14 December 2017

- (9) His **father**, a shopkeeper, died suddenly [...] she still struggles ‘every single day’ with anxiety, having been **sexually abused** as a **child** on a visit **with her family** to Bangladesh and suffered appalling bullying at her primary school.

Times, 30 April 2022

Another cause of ambiguity concerned the potential implications of attending to the uses of the word *sadistic* as a recurrent pre-modifying collocate of abuse. The *Oxford Dictionary of English*⁹ defines *sadistic* as “deriving pleasure from inflicting pain, suffering, or humiliation on others.” This definition is rather vague because it does not specify how “pain, suffering, or humiliation” is inflicted. Similarly, the meaning of *sadistic* is ambiguous in these articles because it might refer to all kinds of abuse, including sexual, as in examples (10)-(12):

- (10) The settings and situations are familiar: the Anglo expats in an idyllic retreat in Provence; the cynical drawing room banter of the effortlessly wealthy; the intimated child **abused** by a **sadistic father**.

Times, 26 May 2018

- (11) Montgomery was **abused** by her **mother** Judy Shaughnessy in **sadistic** ways.

Daily Mirror, 12 January 2021

⁹ Online dictionary, latest update in 2023, version 6.3

(12) The book, published in 2014, detailed the **sadistic abuse** he and his older brother Tom endured at the hands of their **father**, who was head forester on a country estate in Angus, Scotland.

Independent, 30 July 2022

Finally, in other cases incestuous sexual abuse is mentioned but is either (a) stated not to have occurred, or (b) is left ambiguous as to whether or not it did occur. In either case, by mentioning such abuse, the articles imply its relevance to the stories in question, as in examples (13)-(15). This might be because the authors of the articles anticipate that readers will otherwise have suspected, or even expected, incestuous abuse to have occurred.

(13) Police say there is **no sign** that the siblings were **sexually abused**. [...] Safety concerns wouldn't explain, however, why the children were given identical haircuts, the girls all wearing long hair parted in the middle and the boys saddled with the same comical mop top sported by their **father**.

Daily Mail, 17 January 2018

(14) “He **never** spoke of **any sexual abuse**, but he spoke of physical and emotional abuse from his father, and how terrified he was of his **father**, even as an adult,” says Robson.

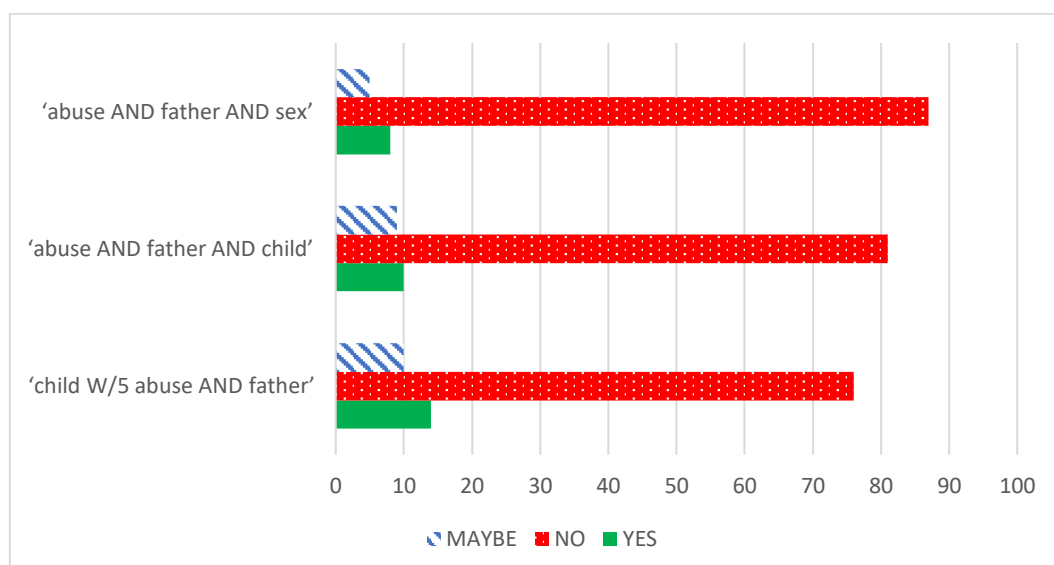
Guardian, 4 March 2019

(15) “I remember my **father** used to come home at 6.30 and my mother used to have his slippers and a glass of sherry ready for him – can this be true?” [...] She suffered **sexual abuse** as a child from an older male, **not a relative**.

Guardian, 16 March 2021

With the different kinds of categorisation in mind, Figure 6.1 shows the results of the close reading of the three English search queries.

Figure 6.1 Results of the close reading of the three English corpus samples



Therefore, most articles in each sample were allocated to the NO category. As observed, the MAYBE category comprises a series of rather complex cases. The vagueness is too strong to confirm that the stories relate to incest, even though they illustrate the ambiguities of the taboo topic. The reasons might be that journalists either decrease the emotional charge of the abuse by being vague or increase it by implying more than what really happened. Thus, the MAYBE category is not included in the corpus. In light of these results, the search query “child W/5 abuse AND father” is considered the most relevant. However, the differences between the three search queries are slight in terms of the proportion of articles classified as being relevant. Thus, the search query needs further filtering to optimise the results.

For this purpose, rather than adding search terms, unwanted meanings of *abuse* are discarded. The aim is to reduce the proportion of articles that could be categorised as being NO or MAYBE, and thereby increase the proportion of articles classifiable as YES. To this end, the function “AND NOT” on Lexis Nexis is used to specify which terms should be excluded. Based on the previous concordance analysis of *abuse* mentioned above, the pre-modifying collocates not referring to *sexual abuse* are excluded from the search query. In doing so, the semantic prosody of *abuse* is narrowed down. Thus, the revised search query is as follows in Table 6.3:

Table 6.3 The final search query for automatically refining the English corpus

| |
|--|
| child W/5 abuse AND father AND NOT alcohol abuse AND NOT clerical abuse AND NOT church abuse AND NOT disability abuse AND NOT drug abuse AND NOT economic abuse AND NOT elder abuse AND NOT financial abuse AND NOT online abuse AND NOT police abuse AND NOT racist abuse AND NOT racial abuse AND NOT substance abuse AND NOT verbal abuse |
|--|

This third search query narrows the corpus down by more than ten times compared to the first corpus, “abuse AND father”, resulting in the collection of 2,019 articles – compared to 23,015 articles. The close reading of 300 articles in the previous stage revealed that incestuous abuse is covered by the press, but indirectly. This indirectness implies complex linguistic patterns. At this point, it was decided that the search terms were still not sufficient to build a corpus based on absence. Therefore, all articles were read to select the relevant ones manually. The number of articles (2,019) was deemed to be more manageable, and thus more amenable to manual filtering of unwanted results. The same categorisation (YES, NO, MAYBE) as previously described was carried out. After manually filtering out results that were not relevant to the research question, 258 articles were collected. Of these, 23 articles use the term *incest* explicitly, which indicates that the search query can retrieve results in which incest is mentioned directly and indirectly.

Finally, the analysis of these 258 articles revealed that members of the family other than father figures were sometimes mentioned. As well as demonstrating that it is not only fathers who are depicted as committing incestuous abuse in the news coverage, this indicates that the search query, while focusing on coverage of incest involving father figures, is not

necessarily so restrictive in terms of the results it retrieves, as examples (16)-(19) demonstrate:

- (16) What you have been through as a child at the hands of your **mother** is, as you describe it, abuse of all kinds: physical, sexual, emotional and psychological.

Times, 9 September 2018

- (17) That Alex had no knowledge that their **mother**, Jill Dudley, had sexually abused them was something Marcus considered a blessing.

Daily Telegraph, 12 October 2019

- (18) A teenage relative began preying on him and for nine-months, he would rape and molest him at every chance he got. "I thought that was my big **brother**, now becomes my rapist."

Guardian, 12 November 2020

- (19) She writes how, aged eight, she watched her maternal **grandfather** sexually abuse her five-year-old sister Kelly, while their grandmother stood at the door to prevent them from escaping.

Sun, 1 April 2021

Furthermore, these examples demonstrate that incest is clearly depicted as child sexual abuse, even though the term incest is not used in the articles. Notably, example (19) illustrates how other members of the family are aware of the incest abuse but do not act against it. This depiction, while only visible in a small minority of the articles in the data, confirms what is described in literature when incest involves the silence of other family members (Dussy, 2021).

2.2. The filtering process in the French corpus

As mentioned above, the iterative approach of filtering the search query had to be adapted to each language due to differences in the semantic prosody of *abuse*: *abus*. Since the meaning of *abus* is more limited in French, the second step of narrowing the search query consisted of excluding the terms pertaining to sexual abuse in the Catholic Church. Thus, the terms *church*, *priest*, *bishop*, *diocese*, *Pope*, and *Francis* were excluded, reducing the corpus by half, from 2,796 to 1,387 articles. Nevertheless, the close reading of 100 articles was deemed necessary to further optimise the search query. As with the English corpus, the articles were classified into three categories (YES, NO, and MAYBE). Similarly, an article was considered relevant if it alluded to incest through explicit mentions of sexual abuse committed by a family member.

As such, articles explicitly mentioning incestuous abuse were classified in the YES category, as examples (20) and (21) illustrate:

- (20) Experienced in sordid cases and used to complex hearings, the officers of the juvenial brigade investigate on violence and **abuse** inflicted on children. [...] The file holds in a nutshell: a teenager girl accuses her **father** of rape, which he denies.¹⁰

La Croix, 19 June 2018

- (21) For the victim, absent parents, for Damien, a **father** violent towards his mother and him, the eldest son. Damien won't admit it spontaneously, but he also experienced **sexual abuse** by an **alliance member of the family**.¹¹

Sud Ouest, 24 September 2020

Notably, example (21) draws on the definition of family given in Chapter 2, i.e., that family ties are not exclusive to blood relations but rather include alliances and adoption. Therefore, this article was included in the YES category, though another researcher may have chosen to discard it due to a different definition of family.

Interestingly, some articles explicitly mention incestuous abuse, but were retrieved through the legal terminology of abuse, such as *abuse of weakness*, as in example (22). The incest perpetrator is clearly identified as a family member (here the *brother*) but the incest is expressed with the explicit terms of *sexual assault* and *rape*, not *abuse*. This demonstrates that articles display a linguistic complexity that is hardly identifiable with an automated search query.

- (22) She reveals to her that she endured **sexual assaults and rapes from her brother**, from the age of 5 to 15. Born from a first union, the recipient's own daughter would also have been a victim of her **stepfather**. [...] The psychological examination [...] concluded that "an **abuse of weakness** had, over time, annihilated his will".¹²

Sud Ouest, 12 July 2019

Conversely, other articles can undoubtedly be classified as NO, since they do not report sexual abuse. The vast majority of this category concerns articles mentioning *abuse* in the field of law, referring to legal infractions such as *abuse of weakness* or *abuse of power*, as mentioned

¹⁰ Agueris aux dossiers sordides et rompus aux auditions complexes, les policiers de la brigade des mineurs enquêtent sur les violences et les **abus** infligés aux enfants. [...] Le dossier tient en quelques mots : une adolescente accuse son **père** de **viol**, ce qu'il nie. (Unless otherwise stated, the French results were translated by this author into English).

¹¹ Pour la victime, des parents absents, pour Damien, un **père** violent envers sa mère et lui, l'aîné des fils. Ce dernier, s'il ne l'évoquera pas spontanément, a aussi fait part d'un épisode **d'abus sexuels** par un **membre rapporté de la famille**.

¹² Cette dernière lui révèle avoir subi des **agressions sexuelles et des viols** de la part de son **frère**, de l'âge de 5 à 15 ans. Née d'un premier lit, la propre fille de la destinataire aurait, elle aussi, été victime de son beau-**père**. [...] L'examen psychologique [...] conclut à « un **abus de faiblesse** qui, au fil du temps, a annihilé sa volonté ».

earlier. Interestingly, some articles write the legal terms between quotation marks as distancing devices, as seen in examples (23) and (24):

- (23) Anthony Ghosn, the French executive's son, stated that his **father** was "ready to defend himself vigorously". [...] This time, the prosecutor's office accuses him of "aggravated **breach [abuse] of trust**".¹³

Les Échos, 7 January 2019

- (24) According to our information, a judicial investigation for "**abuse of weakness**" was opened in August by the Paris prosecutor's office. [...] It notably accuses the lawyer for not respecting the deceased's wish to be buried in **Père-Lachaise** cemetery.¹⁴

Libération, 30 October 2021

Another reason to classify some articles in the NO category is that sexual violence occurs but outside the family circle, as observed earlier with the articles referring to sexual abuse in the Catholic Church. Furthermore, examples (25) and (26) illustrate how the notion of sexual violence is implied in the use of abuse but remains vague and cannot be ascertained. Example (25) is in the context of the #MeToo movement in Greece and example (26) takes place in Afghanistan. This demonstrates that the vagueness of abuse is not only a feature of the English language but can also be perceived in the French corpus.

- (25) Several articles report that he was stabbed by the **father** of one of his victims, but once again, the story was buried. [...] Among the multitude of incidents mentioned in the press, there are **abuses** against young refugees to whom Dimitris Lignadis was giving theatre class.¹⁵

Libération, 25 February 2021

- (26) The **father** gets angry, threatens to inform the top Taliban chiefdom about the **abuses** committed by their men.¹⁶

Le Figaro, 2 September 2021

¹³ Anthony Ghosn, le fils du patron français, a affirmé que son **père** était « prêt à se défendre de façon très vigoureuse ». [...] Cette fois, le parquet l'accuse « **d'abus de confiance** aggravé ».

¹⁴ Selon nos informations, une instruction judiciaire pour « **abus de faiblesse** » a été ouverte en août par le parquet de Paris. [...] Elle reproche notamment à l'avocat de ne pas avoir respecté la volonté du défunt d'être inhumé au **Père-Lachaise**.

¹⁵ Il aurait même été une fois poignardé par le **père** d'une de ses victimes, rapportent plusieurs articles, mais là encore, l'histoire fut enterrée. [...] Parmi la multitude de faits évoqués dans la presse, des **abus** contre de jeunes réfugiés auxquels Dimitris Lignadis donnait des cours de théâtre sont évoqués.

¹⁶ Le **père** se fâche, menace d'avertir la haute chefferie talibane des **abus** que commettent leurs hommes.

Finally, out of 100 articles, only one article may be classified into the MAYBE category as close sentences mention family members and abuse, as seen in example (27). It is worth noting that fewer articles are classified as MAYBE in the French corpus than in the English one. The vagueness around the term *abuse* seems more characteristic in the English corpus, with the use of adjective like *sadistic*, that are absent in the French corpus.

(27) His **father** dies when he is three months old, and he is entrusted to **his mother's cousin** living in Sablé-sur-Sarthe, Berthe Laurens, who adopts him. He traced back to his childhood in a film he directed in 2009, *Sablé-sur-Sarthe, Sarthe*, where he delicately mentions an **abuse** of which he was a victim.¹⁷

L'Humanité, 5 January 2018

As such, over a sample of 100 articles, 85 of them were classified in the NO category, 14 in the YES category, and 1 in the MAYBE category. Therefore, following those observations, the decision was to exclude additional terms to refine the search query. For this reason, the terms related to the legal terminology were excluded as they constituted most of the NO category. Even though there were cases in the YES category also mentioning *breach [lit. abuse] of trust* or *abuse of weakness*, it was decided that the gain of excluding them was greater than keeping them, even though some wanted matches would be missed. As such, the final search query is the following:

Table 6.4 The final search query for automatically refining the French corpus

| |
|--|
| père AND abus AND NOT église AND NOT prêtre AND NOT évêque AND NOT diocèse AND NOT pape AND NOT françois AND NOT abus de faiblesse AND NOT abus de confiance AND NOT abus de biens sociaux AND NOT abus de pouvoir AND NOT abus d'alcool |
| father AND abuse AND NOT church AND NOT priest AND NOT bishop AND NOT diocese AND NOT pope AND NOT françois AND NOT abuse of weakness AND NOT abuse of trust AND NOT abuse of assets AND NOT power abuse AND NOT alcohol abuse |

This third search query retrieved 785 articles, amounting to 553,702 words. This final search query reduced over three times the number of articles compared to the first search query “abus AND père”, which had amounted to 2,796 articles. Thus, this third search query is final as it would be difficult to further include or discard relevant search terms. As observed, articles show linguistic complexity necessitating close reading and manual interpretation. Therefore, and following the iterative approach applied to the English corpus, the remaining 785 articles were thoroughly read to determine if they could be classified in the YES, NO, or MAYBE categories. From this close reading, 223 articles were deemed relevant to the search question.

¹⁷ Son **père** meurt lorsqu'il a trois mois, et il est confié à **une cousine de sa mère** vivant à Sablé-sur-Sarthe, Berthe Laurens, qui l'adopte. Il reviendra sur son enfance dans un film qu'il réalise en 2009, *Sablé-sur-Sarthe, Sarthe*, où il évoque avec pudeur un **abus** dont il fut victime.

Therefore, within the timeframe starting on 15 October 2017 and ending on 14 October 2022, 223 French articles and 258 English articles were identified as relevant according to the set criteria. Both in the French and English corpora, the retrieved articles cover incest as child sexual abuse occurring within the family. However, although the search query aimed to retrieve articles mentioning incestuous stories without directly using the term *incest*, the final French corpus comprises 85 articles explicitly using the term *incest*. In contrast, 23 articles explicitly use the term *incest* in the English corpus, which means that this explicit mention is three times as frequent in the French corpus than in the English one. This indicates that the French corpus assimilates the terms of *abuse* and *incest* more often than the English corpus. Thus, these findings confirm the results observed in the corpora built with the lemma *incest** in Chapter 5, i.e., the French press primarily covers incestuous abuse. However, the English corpus “abuse AND father” shows that incest is also covered as sexual abuse occurring within the family, between an adult and an unconsenting child. Incestuous abuse is covered, and not eroticised, but the term *incest* seems not to apply to this kind of stories predominantly. It further suggests that the understanding of *incest* in English differs from French (see Chapters 2 and 5).

3. Analysing taboo language through signalling nouns

The French and English corpora “abuse AND father” were built on absence, using an iterative approach. The French corpus consists of 223 articles and the English corpus of 258 articles. Each article was thoroughly read to observe direct or indirect mentions of incest. For this reason, these corpora are relevant to investigating taboo language, as the relevant size of the corpus depends on the research question (McEnery & Brookes, 2022). In this thesis, taboo language is defined as linguistic avoidance strategies, generated by the sensitivity of a taboo topic, often silenced. As such, the analysis focuses on discursive absence. Therefore, linguistic taboos may increase vagueness as language becomes purposefully imprecise (Cutting, 2007). One linguistic item that is considered inherently vague is a signalling noun. Signalling Nouns (henceforth SNs) are abstract nouns that encapsulate the content of the preceding or following clauses (Flowerdew & Forest, 2015). As such, SNs are inherently vague, but become specific depending on the context in which they are used. They have been extensively studied in academic writing, but their signalling role in media discourse remains underexplored. Therefore, this section aims to investigate whether signalling nouns contribute to taboo language.

3.1. Corpus linguistics and signalling nouns

SNs are abstract nouns “devoid of specific meaning when taken out of context, but made specific in meaning by their context” (Flowerdew, 2002, p. 147). Words such as *fact*, *idea*, *problem*, *possibility* or *argument* are examples of SNs (Flowerdew, 2002; Flowerdew & Forest, 2015; Schmid, 2012). Indeed, they are vague enough to encompass different meanings depending on their following or preceding clause(s). SNs are called *cataphoric* when they refer to the following clause, and *anaphoric* when they refer to the preceding clause. Therefore,

SNs create lexical cohesion by encapsulating the content of the corresponding clause (Flowerdew, 2002; Jiang & Hyland, 2015; Schmid, 2012). Furthermore, the meaning of SNs relies on “lexical realisation” or “lexical specification” (Flowerdew & Forest, 2015), i.e., SNs complete the meaning of the referring clause, while the referring clause specifies its SN.

In this section, the term *signalling noun*, coined by Flowerdew (2002), is used as it emphasises its discursive function. SNs serve as signposts “making explicit how particular clauses are meant to be interpreted” (Flowerdew & Forest, 2015, p. 34). In other words, SNs signal the “organisation of the discourse” (Ibid., p. 40) by labelling content with a general yet chosen notion. As the focus is on the discursive influence of SNs, Flowerdew’s term was selected. However, it should be mentioned that such a concept has also been labelled *shell noun* (Schmid, 2012), *general noun* (Halliday & Hasan, 1976), *carrier noun* (Ivanič, 1991), *anaphoric noun*, abbreviated *A-noun*, (Francis, 2002), and *stance noun* (Jiang & Hyland, 2015). Interestingly, these terms emphasise different characteristics of SNs. While Ivanič (1991) and Schmid (2012) highlight the lexical function of SNs, Halliday and Hasan (1976), and Francis (2002) emphasise their grammatical features. Finally, Jiang and Hyland (2015) focus on the rhetorical function of SNs when writers employ them in academic writing to indicate their stance on a topic. The term *stance* enhances the metalinguistic function of SNs used to express writers’ perceptions and judgments, especially in academic style (Charles, 2003).

In addition to the varied terminology, SNs have been broken down into different semantic and grammatical categories according to scholars. For instance, Schmid (2012) categorises SNs – or in his case, *shell nouns*– according to their semantic features. SNs can be factual (facts), mental (ideas), linguistic (utterances), modal (possibilities/permission), or eventive (activities). Furthermore, he distinguishes *core shell nouns* from *peripheral shell nouns*, namely SNs that are clearly perceived as abstract nouns from those that have more debatable meanings. SNs such as *fact, feeling, idea, intention, hope, and view* are abstract and unspecific whereas *area, region, and site* are abstract locations that may have physical existence (Schmid, 2012). Therefore, the degree of abstractness varies and must be evaluated when considering whether a term is a SN. Similarly, Flowerdew and Forest (2015) distinguish *constant vs. variable* SNs, i.e., SNs that are context-independent from those that are context-dependent. The question of “whether groups of words are included or excluded from membership in the class of SNs” (p. 40) is crucial for conducting discourse analysis.

Besides these semantic considerations, the grammatical features are critical for analysing SNs. SNs have typical grammatical features such as a noun followed by *that-, wh-, to-*clauses, or clauses with a preposition. These grammatical patterns can be marked as *N-that, N-wh, N-to,* and *N-prep* (Flowerdew & Forest, 2015) or as *N-cl* and *N-be-cl* (Schmid, 2012). There are recognisable grammatical patterns when a signalling noun is introduced. Consequently, SNs can be recognised automatically because they follow identical patterns. For this purpose, computational tools developed in corpus linguistics are extremely helpful for conducting such analysis. Corpus linguistics relies on software to collect large amounts of data and compare data frequencies (Hunston, 2006). Frequencies help to determine common grammatical

patterns whereas infrequent phenomena can be considered as linguistic exceptions. An additional advantage is that observations rely on authentic examples, rather than invented or intuitive ones (Hunston & Francis, 2000; Schmid, 2012). Thus, SNs can be searched through automatic queries that target specific grammatical structures. In the corpus “abuse AND father,” two grammatical patterns were selected. The SNs should be either followed by a clause introduced with the conjunction *that* or the SNs should be introduced by the deictic *this*. These two grammatical patterns are widely investigated (Flowerdew & Forest, 2015; Hunston & Francis, 2000) and are selected for this analysis.

However, one dilemma encountered due to cross-linguistic challenges is that the literature on SNs focuses on the English language, and studies on other languages were not found. Furthermore, the corpus linguistic tools developed to automatically retrieve SNs are based on the grammatical and semantic features of English. Some attempts were conducted on the French corpus but yielded limited results. For instance, the equivalent of the *that*-clause could be the relative clause introduced by *que* (that). Example (28) is one rare finding extracted from the French corpus.

- (28) “I was **abused by my stepfather from 7 to 9 years old** and placed in the Jean-Bru Home on suspicion. I didn’t know that what he was doing to me was wrong. It was only at the age of 10 that I revealed what had happened to me. I had been here for a year, and it was through discussions with other young girls that I understood I had undergone the same **thing as [lit. that]** them”, tells this former resident.¹⁸

(*Le Figaro*, 29 December 2020)

However, it was found challenging to apply the same framework of SNs to the French corpus due to the lack of literature on this topic and the difficulty to develop equivalent corpus linguistic search queries. Furthermore, as example (28) demonstrates, the French findings also need to be translated into English, requiring a comparable transcription of grammatical patterns to English, which complicates the results. Therefore, it was decided to limit this section to the English language.

Thus, this section focuses on the SNs in the English corpus. For this purpose, the corpus “abuse AND father” was uploaded to the software Sketch Engine. The grammatical patterns analysed in the corpus are written as follows:

- (a) [tag= N*] + [word= that]
- (b) [word= this] + [tag= N*]

¹⁸ « J’ai été **abusée par mon beau-père de 7 ans à 9 ans** et placée à la maison Jean-Bru sur suspicion. Je ne savais pas que ce qu’il me faisait était mal. C’est seulement à l’âge de 10 ans que j’ai révélé ce qui m’était arrivé. Cela faisait un an que j’étais ici et c’est au travers de discussions avec d’autres jeunes filles que j’ai compris que j’avais subi la même **chose qu’elles** », raconte cette ancienne.

The SN *knowledge* has a cataphoric function in referring to incest mentioned in the following clause. Furthermore, the explicit mention of rape leaves no doubt regarding the description of incest in this sentence, which corresponds to child sexual abuse. However, mentions of incest were sometimes more circumvented, as in example (31):

- (31) **A terrible, unjust mess that** could have been avoided had all those whose profession centres around protecting innocents such as Poppi done their job properly.
(*Sun*, 17 January 2018)

This example refers to a case reported several times in the press where Poppi, a five-year-old child, died after being raped by her father, who was later convicted. The SN *mess* refers to this tragic event. However, determining the reference of *mess* requires background knowledge. The description of events is scattered in the newspaper article, or sometimes barely mentioned in articles, referring to events already known and covered. It is what Flowerdew (2002) refers to as an *exophoric reference*, i.e., SNs understood thanks to the “reader’s or listener’s background knowledge” (p. 150). However, in the analysis, only *endophoric references* of SNs are analysed, i.e., the reference is within the text, because the focus is on how the referring clauses are signalled. Thus, whenever the reference can be clearly established in the corresponding clause, they are included in the analysis.

Criterion (b) addresses the degree of abstractness concerning SNs. Terms such as *belief*, *case*, *claim*, *idea*, *fact*, and *thing*, are considered prime examples of abstract nouns in the way they conceptualise vague notions (Flowerdew & Forest, 2015; Jiang & Hyland, 2015; Schmid, 2012). However, a term such as *secret* seems more questionable. A secret seems more specific than a *fact* because it carries more implications (e.g., silence, confidence, hiding), yet it still requires content to be fully defined, as in example (32).

- (32) Connolly has previously described **the sexual abuse** he suffered as a **dark secret that** he had shouldered throughout adulthood.
(*Times*, 21 October 2019)

Similarly, Schmid classifies *boy* as a general noun, and more specifically as “human noun,” along with *people*, *person*, *man*, *woman* and *girl* (2012, p. 274). However, this classification is questionable in example (33):

- (33) This continued for a period of time and the defendant threatened **the boy that** if he told anyone he would **sexually assault his sister**.
(*Independent*, 21 November 2017)

The *boy* refers here to a brother. Although the noun *boy* may be considered abstract, the context in which it is used makes it very specific. In this example, the noun *boy* serves as a substitute for the noun *brother*, but *boy* is specific rather than vague or abstract. It is not any boy that is discussed, but a boy perceived through the kinship with *his sister*. Hence, the inclusion or exclusion of those SNs tends to precede the analysis of what constitutes a SN and what it refers to. Following these considerations, out of 1,035 occurrences of *N-that*, 142 were

considered relevant to the criteria established to answer the research question, and out of 73 occurrences of *This+N*, 11 were considered relevant. In this corpus, SNs serve four main functions: a substituting function, in 61 cases; a renaming function, in 56 cases; a timing function; in 19 cases; and, more rarely, a defining function; in 6 cases. The following subsection explains these categories.

3.2. A taxonomy of signalling nouns on a taboo

In total, 153 SNs following the patterns *N-that* and *This+N* correspond to the criteria mentioned above: they refer to incest and are abstract nouns. Out of these 142 SNs, 61 were classified in the substituting category, 56 in the renaming category, 19 in the timing category, and 6 in the defining category. The function of substitution is analysed first, and the other functions are explained further on.

Substituting

In the corpus, 61 SNs contribute to increasing the vagueness of the coverage of incest. It was previously noted that SNs and their referring clause are complementary for realising lexical specification. In other words, SNs complete the meaning of the clause and are completed by the referring clause. However, examples (34)-(38) illustrate how both the meaning of the clause and its corresponding SN are vague:

- (34) We don't think about what's driving people towards drugs; we might think about regulating access, when actually it's the consequences of **something that happened to someone as a child.** "Abused children often become hyper-vigilant, Bellis explains.

(*Guardian*, 28 June 2018)

- (35) The statement said that she was still uncovering the ways **the abuse** has affected her life and said she had been hurt "**in a way that** no human should".

(*Independent*, 14 May 2022)

- (36) **This kind of abuse** warps something inside of you. Because it doesn't happen by a stranger that snatches you off the street; **it happens by someone you love and trust.**

(*Times*, 13 March 2021)

In examples (34) and (35), the abstract nouns *something* and *way* are SNs as they are inherently vague (Flowerdew & Forest, 2015; Schmid, 2012). In this context, they reinforce the vagueness of the clause as they do not provide any additional meaning. In these two examples, the term *abuse* is not explicitly characterised. To some extent, *abuse* can be considered as a SN for its inherent vagueness. Additionally, example (9) shows how *abuse* is further unspecified through the vague classification *[t]his kind of*. Thus, SN does not correspond to the definition of being unspecific when isolated and becoming specific when in

context (Flowerdew & Forest, 2015; Schmid, 2012). On the contrary, SNs remain unspecific and euphemise the taboo further. As such, SNs serve as substitutes for more specific words.

Renaming

Conversely, there are specific mentions of incestuous abuse in the clause, but SNs keep at a distance what is reported in the newspapers. In this case, SNs rename the referring clause. For instance, the term *allegation* throws doubts to what is reported. Examples (37) and (38) give explicit descriptions of incestuous abuse. In both cases, it involves the father sexually assaulting an underage (step)daughter. However, these descriptions are distanced through the SN *allegation*. This term often occurs in media discourse to cover cases of sexual abuse because journalists report information while protecting themselves against defamation (Tranchese, 2023). Nevertheless, the term tends to protect the abusers over the victims, even in situations where there is no risk of defamation (Ibid.). It is not a mere description; it is an alleged version made by the victims. For this reason, the victims' testimonies are undermined. Therefore, this type of SN was considered to rename the clause. Instead of specifying it, it tends to undermine it by encapsulating the description in a loaded term.

(37) Most of our younger siblings had long since gone into care, following **allegations that my father had raped my sister** Louise, 13.

(*Mail*, 3 September 2018)

(38) He also believed **the mother's allegation that** several years earlier she had seen **the father with an erection, leaning over and touching his five-year-old stepdaughter** while her pyjama pants were down, [...].

(*Guardian*, 19 September 2019)

Furthermore, SNs can influence the discourse in the way they determine the meaning of the clause. For instance, examples (39) and (40) show that for the same SN, the inferred meaning differs according to the highlighted and chosen perspective. Indeed, the SN *claim* either confirms or invalidates the victim's testimony. In example (39), the *claim* is the stepfather's, asserting that he was asleep during the sexual act. The possessive *his* shows that his version is chosen over the one of the *teenager*. However, in example (40), *claim* takes the perspective of the victim. Firstly, it is mentioned through the possessive pronoun *her father*, emphasising the chosen perspective. Secondly, the claim is contested through the direct quotation where she reiterates her claim that she has been *sexually abused by [her] father*. Therefore, the father's claim is further weakened through the daughter's testimony within the sentence. Thus, SNs realise the lexical cohesion and organise how the discourse should be interpreted.

(39) The Pamplona court accepted **his claim that** the teenager had **sat astride him** while he was asleep on the sofa and had **engaged in penetrative sex**.

(*Telegraph*, 16 April 2021)

- (40) Asked about **her father's claim that** she had been coached, she replied: "What I don't understand is how is this crazy story of me being brainwashed and coached more believable than what I'm saying about **being sexually assaulted by my father**?"
(*Times*, 19 January 2018)

In the previous examples, the renaming function is cataphoric because SNs refer to the following clause. However, the renaming function can be anaphoric, as in examples (41) and (42). Interestingly, in example (41), the SN *trauma* encapsulates the preceding clause, renaming the incestuous abuse. In this case specifically, the SN adds information by enhancing the lasting consequence of incest. Conversely, in example (42), *[this] abuse* refers to an explicit incestuous abuse committed by the father, but renamed vaguely in the following sentence.

- (41) **Paul's adoptive mother sexually assaulted him from the age of three, a trauma that** has left him unable as an adult to have intimate relationships with women.

(*Independent*, 21 July 2021)

- (42) [...] what he once called in an interview the "big rucksack full of bricks" that constituted his memories of **his father's sexual abuse of him as a child**. Connolly talks to me a little about this, of his unnecessary guilt, of how he felt too much love for his father to feel angry, but really does not want to talk about it. **This abuse** does not feature, either, in his new book, *Made in Scotland*, a memoir of his early life in Glasgow.
(*Times*, 29 November 2018)

Therefore, the function of renaming contributes to lexical realisation, as seen with the SNs *trauma* and *claim*. However, it also contributes to vagueness by undermining the meaning of the clause, as seen with the SNs *allegations*. Thus, SNs shape "the organisation of discourse" (Flowerdew & Forest, 2015, p. 40) by influencing how discourse should be interpreted.

Timing

Schmid (2012) defines the SNs of time as circumstantial. Flowerdew and Forest (2015) offer a similar definition, and consider time, along with place, as *circumstantial facts*, because they "present information in terms of where, when, how, and why" (p. 32). Notably, there are many SNs referring to time in the corpus. The particularity of incest as child sexual abuse is that children are often helpless to report the assault when it occurs, and finally reveal the facts once they are adults (CIIVISE, 2023). Therefore, time is a key element in the narration of incest, as in examples (43) and (44).

- (43) **This week**, in her first TV interview about her shocking allegations, Dylan **recalled how**, while her mother was out, Allen took her to an attic space.

(*Mail*, 19 January 2018)

- (44) Everyone loves him, everyone appreciates him, and in fact no one imagines **for a single moment that** he is **his stepdaughter's lover**.

(*Times*, 19 August 2019)

In examples (43) and (44), the SNs *week* and *moment* mark vague temporality. In example (43), the interplay between time (*week*) and testimony (*recall*) shows that temporality is central to the discourse on incest. Drawing a parallel with Schmid's (2012) analysis of the SN *journey* – it is challenging to know when a *journey* starts (from home or once on the train?) and where it ends –, a week may refer to different time periods and convey different meanings, depending on whether it is a *working week* or a *holiday week*. In this case, *this week* is a critical temporal marker in the narration. In example (44) signals vague and unfathomable temporality. The references to time highlight the negation of incest through disbelief or rejection: *no one imagines for a single moment*. Therefore, time either marks evolving temporality in the discourse of incest, or enhances the impossibility of its occurrence.

Defining

In rare occasions, 6 out of 153 relevant occurrences, SNs enable explicit lexical realisation. In other words, SNs contribute to defining more specifically the referring clause. Example (45) shows a rare occurrence of the word *incest* in the corpus. Interestingly, the classification [*t*]*he particular kind of abuse*, which is unspecific, is made specific through the incision of *incest* in the following clause. Similarly, example (46) shows another cataphoric reference of the SN *something*. In this case, the lexical specification through the referring clause introduced by the colon. The scene that *disturbed [the babysitter]* is explicitly stated; there is no vagueness in the description, even though the scene is not *stricto sensu* depicting incestuous assault.

(45) **The particular kind of abuse that Dylan alleges against Allen, incest**, also has recognizable patterns.

(*Guardian*, 13 March 2021)

(46) Her babysitter told her she'd seen **something that disturbed her: Dylan on the sofa, with Allen's head in her lap.**

(*Times*, 12 March 2021)

Furthermore, SNs have a defining function in the way they are used metalinguistically. SNs serve to comment on the terms chosen to report incest. The metalinguistic comments can clarify or increase the vagueness of the terms. In example (47), the clause *incest is what happens* adds vague information on incest. Conversely, in example (48), the victim considers the term *incest* as a euphemism for *child rape*. As such, the SN *thing* introduces his metalinguistic claim in favour of a different use of terms.

(47) "**Incest**" is a **term that** often implies consent, or at least some sort of mutuality, but in reality, **incest is what happens when** the power imbalances inherent in the family structure are harnessed as **a tool of sexual assault**.

(*Guardian*, 13 March 2021)

(48) "The **other thing that** bothers me, I will say this," he explains, "is the sanitisation of words like '**incest**' and '**sex abuse**,' instead of '**child rape**.'"

(*Telegraph*, 27 November 2021)

Therefore, SNs serve several discursive functions regarding the taboo topic of incest: substituting, renaming, timing, and defining. As observed, SNs tend to obscure the lexical realisation by maintaining vagueness. Lexical realisation refers to the mutual specification of the SN and the clause (Flowerdew & Forest, 2015; Schmid, 2012). The meaning of the clause is encapsulated in the SN, giving a meaning to the abstract SN. However, this definition is valid in academic writing where SNs have been extensively analysed (Charles, 2003; Flowerdew & Forest, 2015; Jiang & Hyland, 2015; Schmid, 2012), but less relevant in media discourse, especially when a taboo topic is reported. Therefore, SNs may be used precisely because of their inherent abstractness. In other words, SNs are the perfect tools for euphemising a taboo topic. Inherently, SN “function as containers,” thus, “it is necessary that they have a gap, a hole, or some other kind of opening or dent which can receive the content” (Schmid, 2012, p. 76). It is perhaps where lies the paradox for this analysis: SNs are the perfect terms to euphemise taboo topics.

4. Summary

As the analysis above shows, incest can be directly or indirectly referred to. This chapter raises methodological challenges for collecting such implicit occurrences and attempts to address them with a different corpus linguistic approach. Furthermore, the notion of SNs provides an additional layer to understand taboo language on incest. Using SNs in discourse “affects the reader’s perception” and “enables the reader to perceive the organisation and meaning that the writer intends” (Charles, 2003, p. 318). SNs are abstract yet chosen nouns that structure discourse and influence its understanding.

Finally, these avoidance strategies for naming the taboo raise broader considerations on how sexual violence is discussed, or “labelled.” As Vessey (2024) points out, “terminological vagueness” tends to categorise discourse on sexual violence. However, “labels for sexual violence” matter (Ibid., pp. 2-3):

They matter because they affect the extent to which acts are seen as being inappropriate or violent, much less the extent to which they are understood to have happened at all. Labels that are too broad and all-encompassing can be vague and lose their utility; labels that are highly precise can become niche and specialized; all labels may have specific denotative, connotative, and legal entailments that vary across different sites.

Thus, analysing terminological vagueness on incest poses the larger question of how taboo topics are framed in the press. To further investigate this notion of taboo language, the next chapter focuses on a case study dedicated to the media coverage of Woody Allen found both in the corpora built with the lemma *incest** and the search query “abuse AND father,” and both in the French and English press. This case study investigates how linguistic choices discursively construct taboo in relation to a celebrity.

Chapter 7: Taboo and Celebrity, a case study on Woody Allen and Incest

Celebrity culture revolves around rumour and hearsay as much as facts and professional reporting.

(Rojek, 2012, p. 7)

This chapter delves into a more fine-grained critical analysis on the incest taboo in relation with a celebrity. It was noticed during the previous analyses that Woody Allen's name occurs in the French and English corpora built through the lemma *incest** and through the search query "abuse AND father" (see Chapters 5 and 6). Notably, Woody Allen was the *only* name occurring across the French and British press. Indeed, in 1992, his seven-year-old adoptive daughter, Dylan Farrow, accused him of incestuous abuse. For this reason, it is worth examining how the incest taboo and this American celebrity, well-known both in France and the UK, are discussed in the press. Thus, this chapter aims to investigate the connection between a taboo topic and a celebrity cross-linguistically and cross-culturally since "celebrity culture is characterized by globalization as actors, politicians, and reality stars reach celebrity status worldwide" (Van Den Bulck et al., 2017, p. 48). For this purpose, the subcorpus "Woody Allen" was built by collecting articles through the search term "Woody Allen" in both corpora. In total, 50 articles were retrieved from the British press (36 from the corpus "abuse AND father" and 20 from the corpus *incest**), and 14 articles were retrieved from the French press (9 from the corpus "abus AND père" and 7 from the corpus *incest**). Firstly, this chapter introduces the characteristics of celebrity news. Then, the portrayals of Woody Allen as "rogue celebrity" and Dylan Farrow, as the accuser and victim, are analysed. Finally, the taboo language framing this taboo topic is examined to shed light on the various linguistic avoidance strategies.

1. Celebrity news

Celebrities are famous individuals, recognised for their deeds and scrutinised in the media. Usually, celebrities "have emerged from the sports or entertainment industries, they will be highly visible through the media, and their private life will attract greater public interest than their professional life" (Turner, 2014b, p. 3). Three types of celebrities can be differentiated according to Rojek (2001): ascribed, achieved and attributed. The ascribed celebrity is famous by birth, such as royal family members or celebrities' children who inherit fame. The achieved celebrity becomes famous through their accomplishments, such as athletes or actors. The attributed celebrity attracts attention through scandals, including sex scandals or crimes.

Following this categorisation, Woody Allen can be characterised as an achieved celebrity, since he became famous due to his filmmaking and acting. Indeed, Woody Allen is the subject of several books that analyse his movies and personality, presenting him, for example, as “one of America’s most prolific filmmakers, [...], his work is highly recognizable for its seemingly autobiographical themes and New York settings” (Allen et al., 2006, p. vii). He is also an actor, often playing in the films he directs, such as *Take the money and run* (1969), *Everything you always wanted to know about sex (but were afraid to ask)* (1972), *Annie Hall* (1977), *Manhattan* (1979), *Husbands and wives* (1992), *Celebrity* (1998), *Match Point* (2005), and many others, as he directs one film per year on average. His 50th and latest film, *Coup de Chance* (lit. “stroke of luck”) in 2024, takes place in France. These recent years, Woody Allen primarily set his camera in Europe to narrate new stories, as he “has enjoyed a better relationship with European than American critics” (Allen et al., 2006, p. ix). Thus, in addition to being an American celebrity, he also enjoys fame in Europe, which is why he receives coverage in both the French and British press.

Thus, Woody Allen is a celebrity, participating in the celebrity culture that gained momentum with the expansion of mass media after the Second World War (Rojek, 2001). Celebrity culture is constructed through the mass media and contributes to the thriving of the mass media. As such, “celebrity is a genre of representation and a discursive effect” (Turner, 2014b, p. 10). This is why investigating celebrity news reveals social phenomena (Rojek, 2001, 2012; Turner, 2014b). Celebrity news refers to the media coverage focusing on celebrities (Dubied & Hanitzsch, 2014; Van Den Bulck et al., 2017). The increasing media coverage about celebrities suggests that celebrity news is an integral feature of today’s journalism (Turner, 1999), even though celebrity news is criticised for undermining serious or “quality” journalism and relaying sensational events (O’Neill, 2012). Thus, the term *tabloidisation* is often used to describe the “dumbing down” of the quality press, i.e., the broadsheets, that increasingly devote larger space to “light” or “soft” news, such as celebrity news, instead of reporting “hard” news, such as political and economic issues (Gripsrud, 2000; Turner, 1999). In other words, the evaluation of what is deemed newsworthy, i.e., newsworthiness, has shifted: “the intimate relationships of celebrities from soap operas, the world of sport or the royal family are judged more ‘newsworthy’ than the reporting of significant issues and events of international consequence” (Franklin, 1997, p. 4). Despite referring initially to the printed press, the term *tabloidisation* became an umbrella term to pinpoint the spreading of tabloids’ characteristics to all forms of media, including television. Hence, *tabloidisation* “moves beyond the description of a particular kind of journalism to become a portmanteau description for what is regarded as the trivialisation of media content in general” (Turner, 2014b, p. 84). Notably, the term *tabloidisation* is common in English-speaking literature, and significant regarding the British press, as tabloids represent the most circulated newspapers in the UK (McNair, 2009; Office of Communications, 2023). However, tabloids do not exist in the French media landscape, as previously noted in Chapter 4. Instead, the term *sensationnalisme* (sensationalism) is used to describe the press’ increasing focus on sensational events (Awad, 1995; Labasse, 2012). Furthermore, the term *sensationalism* is used in English as well, either as a characteristic feature of tabloidisation, as in Franklin (1997):

“measured judgement has succumbed to *sensationalism*” (p. 4, emphasis added), or as an alternative term to *tabloidisation* (Brown et al., 2018). In this chapter, *tabloidisation* and *sensationalism* are equally used to refer to the media’s propensity to cover soft news, like celebrity news, instead of reporting “hard” news, both in France and the UK.

This stated, it is worth mentioning that the debate on sensationalism is divided between those who believe that the report of celebrity news favours disinformation and the “dumbing down” of the press in general (Bourdieu, 1996; Couldry & Markham, 2007; Franklin, 1997; O’Neill, 2012; Turner, 2016), and those who believe that it brings diversity and accessibility to information (Gripsrud, 2000; Johansson, 2006; Keeble, 2009; McNair, 2009). For this reason, Couldry and Markham (2007) conducted research to determine if the coverage of sensational news precludes or favours “democratic engagement”. Their study demonstrates that those who engage the most with celebrity news are less aware of public issues, calling into question the assumption of “democratisation” through the consumption of celebrity news (Turner, 2016). Therefore, celebrity news diverts public opinion from serious topics such as political, economic or social affairs to cover sensational topics often involving sex and drama (Brown et al., 2018; O’Neill, 2012). In doing so, celebrity news may have a political agenda as they “constitute an essentially conservative, or even reactionary, category of news content that tends to preserve existing social inequalities, thus contributing to the conservation of materialist hegemony” (Dubied & Hanitzsch, 2014, p. 38). This is why CDA, which informed the analysis carried out in this chapter, provides an interesting framework to investigate the ideological agenda of celebrity news thriving in capitalist and materialist societies since CDA contributes to investigating the powers at stake in discourse (Fairclough, 2013; Machin & Mayr, 2023; Van Dijk, 2001, 2006; Wodak, 2014, see Chapter 4).

On the other hand, some argue that perceiving tabloidisation as a sign of qualitative “dumbing down” is a rather elitist view since considering soft news more popular and less qualitative and hard news more serious and qualitative relies on value judgements that are far from being neutral (Gripsrud, 2000; Keeble, 2009; Labasse, 2012). Moreover, it tends to consider that the audience consumes only one type of news and cannot rejoice in finding information both in soft and hard news. Furthermore, McNair (McNair, 2009) notes that in the past, when the number of media outlets was more limited and primarily produced by a small number of (male) journalists, it did not offer as many types of news as it does today. The media has diversified over the last years, allowing individuals and communities to find various information. Thus, the phenomenon of tabloidisation in the media, and the press in particular, is not necessarily a “bad” thing (Gripsrud, 2000). On the contrary, sensationalism can be perceived as offering another perspective to information (Gripsrud, 2000; McNair, 2009). For this reason, the coverage of less important news may provide other perspectives on the world that are rarely reported (Keeble, 2009). In this sense, instead of selecting events according to their news value of negativity (Bednarek & Caple, 2019; Harcup & O’Neill, 2001), celebrity news favours positive and light events (Dubied & Hanitzsch, 2014; Van Den Bulck et al., 2017).

Finally, the belief that the press's quality has declined suggests that there was once a golden age of journalism. However, there was never a "golden age" where only "serious" news was reported (Gripsrud, 2000, p. 287). Since the founding of the press in the 17th century, "[s]candal has always been the staple ingredient of the mass media and few periods have escaped moral panics over alleged declining media standards" (Keeble, 2009, p. 159). Indeed, tabloidisation finds its origins in oral folk culture in "its accessibility, its use of vernacular, its vulgarity and emphasis on spectacle and excess, its blatant disregard for formal politics and so on" (Biressi & Nunn, 2008, p. 2).

With all this in mind, the next section introduces a case study of celebrity news on Woody Allen and the incest taboo. It investigates how the portrayal of a celebrity is articulated to a taboo topic. The subsidiary research question is to wonder how the status of celebrity might influence the reporting of a taboo crime. In other words, the press might cover a taboo topic because it becomes sensational, thus more newsworthy, when a celebrity is involved.

2. The rogue celebrity

According to the categories defined by Rojek (2001), Woody Allen can be considered an achieved celebrity, who gained fame in directing films and playing in most of them. However, another label could be added to the one of achieved celebrity: rogue celebrity. Penfold-Mounce (2009) uses this term to define celebrities who lose their good reputation due to their crimes or accusations of crimes. They are "deglamorised", i.e., their admired image is stained. Penfold-Mounce conceptualises the notion of rogue celebrity into three categories: celebrity deviant, celebrity suspect, and celebrity criminal. The celebrity deviant has not necessarily committed a crime but has a "naughty image", such as the figure of the bad boy. The celebrity suspect is "suspected or associated with a crime whether or not the ultimate outcome is proven innocence" (p. 143). Finally, the celebrity criminal is proven guilty of a crime. Considering this, Woody Allen would correspond to the notion of celebrity suspect.

The facts are the following. In 1992, Dylan Farrow, the adopted daughter of the actress Mia Farrow and director Woody Allen, accused her adoptive father of sexually abusing her in the attic of Mia's house in Connecticut. 1992 is also the year when Mia Farrow discovered that Woody Allen has a relationship with her other adoptive daughter, Soon-Yi Previn. The twelve-year relationship between Mia Farrow and Woody Allen ended following these two events. At the time of the accusations, Dylan was seven years old, and due to her alleged fragility, the criminal investigations in New York and Connecticut were dropped. Woody Allen, 56 years old at that time, had by then officialised his relationship with Soon-Yi, 21 years old. They married two years later. Woody Allen received extensive media coverage during this period. However, the accusations resurfaced years later with two open letters written by Dylan Farrow. The first letter was written in 2014 on a journalist's blog of the *New York Times* (Farrow, 2014), as the newspapers refused to publish the letter directly in their papers. The second letter was published in 2018 in the *Los Angeles Times* (Farrow, 2017), titled "Why has the #MeToo revolution spared Woody Allen?". The letter explicitly referred to the #MeToo movement that

began in October 2017 (Kantor & Twohey, 2017). Notably, one of the prominent journalists involved in exposing Harvey Weinstein was Ronan Farrow (Farrow, 2017), Dylan's brother and Woody Allen's only biological son. In January 2018, Dylan Farrow gave an interview to the American TV channel CBS and repeated her accusations. It was the first time she appeared on camera to discuss the matter. On 21 February 2021, the four-part documentary "Allen vs. Farrow" (Dick & Ziering, 2021) is released on HBO. The documentary extensively covers the details of the case. It features Mia and Dylan Farrow, along with family friends and relatives, relating the accusations of sexual abuse for the first time. Woody Allen and his wife Soon-Yi refused to participate. It is worth noting that Woody Allen embodies two types of incest. Dylan Farrow's accusations align with the definition of incest as intrafamilial sexual abuse. Allen's involvement with Soon-Yi, Mia's adoptive daughter, after knowing her for twelve years, may fall under the definition of incest as a desired but transgressive relationship within the family. The latter relationship stirred controversy in the 1990s when it became public (Dick & Ziering, 2021), and the following analysis reveals that their marriage is still perceived as transgressive.

Penfold-Mounce (2009) describes three ways for rogue celebrities to avoid deglamorization, i.e., a bad reputation. The rogue celebrity can apologise for the wrongdoing and express their regrets. It is "penance" as it assimilates with "self-punishment" inspired by Christianity (pp. 152-156). The other response is a public declaration to deny or acknowledge the accusations. The third option is victimhood when celebrities present themselves as the victims of an unfair process. It can be questioned whether public declaration and victimhood can be distinguished, since often the victimisation process of a celebrity is performed through a public discourse. In the case of Woody Allen, when the accusations were made in 1992, he made a speech in front of the hospital where Dylan had been examined, to declare his innocence and present himself as a victim: his ex-partner, Mia Farrow, would seek revenge of a painful breakup through their adoptive daughter (Dick & Ziering, 2021). Therefore, it is worth investigating how the press portrays Woody Allen. The coverage may also choose to portray him as a victim, or as rogue celebrity. As for Dylan, his adoptive daughter and accuser, her portrayal may indicate how the accusations are believed or questioned, as a reverse mirror to Woody Allen. Indeed, if the film director is the victim, Dylan is the liar, but if she is believed, he is the criminal or rogue celebrity.

In the subcorpus on Woody Allen, the film director and his adoptive daughter, Dylan Farrow, are the two main protagonists. Interestingly, the way they are represented in the press indicates that the reputation of the film director was not tarnished by the accusations before 2021 but became more suspicious following the release of the documentary in 2021. Informed by the analytical framework of social actors (Van Leeuwen, 1996), the analysis examines how Woody Allen and Dylan Farrow are portrayed in the French and British press, as their nomination influences how the accusations of incest are discussed in the press. As such, nomination "can reallocate roles, rearrange the social relations between the participants" (Van Leeuwen, 1996, p. 43), as seen in examples (1)-(3) where Woody Allen is primarily portrayed as a prominent celebrity in the film industry whose films are part of his persona.

(1) In Hollywood, **the 84-year-old New York director** remains highly regarded.¹⁹

Figaro, 28 November 2017

(2) It's easy to understand her frustration as other men are named and shamed often by anonymous accusers for alleged offences that, frankly, pale by comparison to what she says **the 82-year-old director of *Manhattan* and *Annie Hall*** did to her as a child.

Daily Mail, 19 January 2018

(3) "Daddy" is the **Oscar-winning director, actor and writer** - and Mia's former partner of 12 years - **Woody Allen**.

Times, 13 March 2021

In addition to being nominated, Woody Allen is functionalised. Functionalisation "occurs when social actors are referred to in terms of an activity, in terms of something they do" (Van Leeuwen, 1996, p. 54). Woody Allen is defined as an artist of the film industry. By contrast, Dylan Farrow is solely nominated. Furthermore, she is described through relational identification. Hence, instead of being described for what she *does*, she is determined by what she *is* and to *whom*. Her identity is described through her adoptive father Woody Allen, as seen in examples (4) and (5).

(4) As **the adopted daughter of the actor and director Woody Allen** pointed out, she has been saying for more than 25 years that he molested her [...].

Daily Mail, 19 January 2018

(5) **Woody Allen and Mia Farrow's adopted daughter, Dylan Farrow**, 35 years old, opens the family album.²⁰

Libération, 26 February 2021

Furthermore, in the framing of the accusations, Woody Allen assumes the active role as the grammatical subject, while Dylan Farrow assumes the passive role as the possessed object (Halliday & Matthiessen, 2004; Van Leeuwen, 1996), as illustrated in examples (6) and (7).

(6) **Mr Allen's daughter, Dylan Farrow**, wrote an open letter in 2014 renewing her claims that he sexually abused her when she was a child in 1992.

Telegraph, 20 October 2018

¹⁹ À Hollywood, **le réalisateur new-yorkais de 84 ans** demeure en odeur de sainteté. (Unless otherwise stated, the French results were translated into English by this author.)

²⁰ **Fille adoptive de Woody Allen et Mia Farrow, Dylan Farrow**, 35 ans, entrouvre l'album de famille.

- (7) Caught up in the #MeToo wave and the accusations of sexual abuse by **his adopted daughter Dylan**, Allen is at a standstill.²¹

Figaro, 5 February 2019

However, these examples were retrieved from articles published prior to the release of the documentary “Allen vs. Farrow” in 2021. Interestingly, there is a drastic change in the subjection after the documentary’s release. Dylan Farrow is no longer passivated but becomes the grammatical subject and *her* father the possessed object, as in examples (8) and (9).

- (8) Here’s what’s been known for years: **Dylan Farrow** alleges **Woody Allen, her adoptive father**, sexually abused her when she was a child.

Independent, 22 February 2021

- (9) She watches these images of herself as a little girl with immense emotion, recounting the molestation that she has never stopped accusing **her adoptive father, filmmaker Woody Allen**.²²

Dauphiné Libéré, 2 March 2021

In addition, Woody Allen’s and Dylan Farrow’s voices suggest the perspective chosen by the journalist. Indeed, “[t]he decisions made at the level of *voice representation* may reflect the way journalists feel about the events reported and the emphasis they decide to place on them” (Piazza, 2009, p. 171, emphasis in original). Interestingly, Woody Allen is primarily quoted for denying the accusations before 2021, i.e., before the release of the documentary, while Dylan Farrow is primarily quoted for reasserting her accusations after 2021. Denial and accusation are framed through their reported speech, as seen in examples (10)-(13).

- (10) As for his own alleged crimes, he vehemently denied them in an op-ed on February 7, 2014, in *The New York Times*, concluding with a warning that it would be his **“final word on the matter.”**²³

Figaro, 28 November 2017

- (11) During the interview Allen said it was illogical that he would choose **“to become a child molester”** during a custody battle.

Times, 19 January 2018

²¹ Rattrapé par la vague #MeToo et les accusations d’abus sexuels de sa fille adoptive Dylan, Allen est à l’arrêt.

²² Elle regarde avec une immense émotion ces images d’elle, petite fille, racontant les attouchements dont elle n’a jamais cessé d’accuser **son père adoptif, le cinéaste Woody Allen**.

²³ Quant à ses propres crimes supposés, il a pu les nier farouchement dans une tribune le 7 février 2014 au *New York Times*, avertissant en conclusion que ce serait là son « **dernier mot sur toute cette affaire** ».

- (12) “He **touched my privates,**” says the little girl. “And then he was breathing on my leg. “He said, ‘What about some father daughter time,’” she **continues,** chewing on a strand of hair.

Times, 13 March 2021

- (13) “He rested his head on my lap, hugged me very tightly, and after, after..., secretly, with his hand, here, he **touched me here.** I didn’t like that at all”.²⁴

Dauphiné libéré, 2 March 2021

In addition, it is worth noting that Woody Allen denies the accusations of being a *child molester*, thus euphemising the accusations of incest. On the other hand, Dylan Farrow’s reported speech is euphemistic, i.e., she recounts the assault through vague terms: “*touched my privates,*” “*touched me here,*” “I didn’t like *that.*” Such avoidance linguistic strategies are euphemistic substitutes for *genitals* or *vagina*. Dylan’s voice is the one of a seven-year-old girl narrating the assault in front of her mother’s camera and made public for the first time in the documentary (Dick & Ziering, 2021). As such, a seven-year-old girl may be aware of this taboo word or just unaware of the proper name. Nevertheless, sex remains one of the most tabooed topics in all languages (Allan, 2019; Allan & Burrige, 2006; Crespo-Fernández, 2015, 2018; Pedraza, 2018, see Chapter 3). For this reason, the journalist’s choice of reporting these testimonies can be seen as sensationalistic because they rely on sexual topics to provoke emotions and capture audience’s interest (D. K. Brown et al., 2018). As such, it may be characteristic of tabloidisation (Turner, 1999; Van Den Bulck et al., 2017). Despite these considerations, Dylan’s voice is at the centre of the articles, as shown by the reporting verbs *continue, say* or *speak*, which was not the case before the release of the documentary in 2021.

Finally, the coverage of the accusations against Woody Allen involves tierce protagonists that are other celebrities. Their opinions are reported and seem to be arguments of authority. Actors, including “celeactors” (Rojek, 2001) who are fictional characters that appear “real” to the public like James Bond, publicly supported Woody Allen. Examples (14) and (15) illustrate how celebrity news mixes “facts and fictions, that is, of the characteristics of factual (news) and fictional (entertainment) genres” (Dubied & Hanitzsch, 2014, p. 46). This blending between the glamorous and the serious allegations distances the accusations from reality.

- (14) **Javier Bardem** has **criticised** the “public lynching” of **Woody Allen** and said he would work with the director again tomorrow. The **Oscar-winning Spanish actor** made the **2008 film *Vicky Cristina Barcelona*** with Allen and said he was aware then of the **child sexual assault allegations**, which have resurfaced in recent years.

Telegraph, 20 October 2018

²⁴ « Il a posé sa tête sur mes genoux, m'a serré très fort dans ses bras, et après, après..., secrètement, avec sa main, ici, il m'a touché ici. Je n'ai pas aimé ça du tout ».

- (15) **From Scarlett Johansson to Pierce Brosnan**, including **Catherine Deneuve**, **personalities** have **come to the defense** of Woody Allen, who has been boycotted by distributors in the United States due to **accusations of sexual abuse**. This occurred during the screening of his **new film *A Rainy Day in New York*** (released in France on September 18) at the Deauville Festival. **“Woody Allen is in the Pantheon of directors. His work is meaningful,” declared former 007 Pierce Brosnan.**²⁵

Sud Ouest, 7 September 2019

Nevertheless, it is worth noting that some actors distanced themselves and withdrew their support after #MeToo movement in 2017, reinforced after the documentary in 2021, as seen in examples, (16) and (18). We can note a gradation in the precision of the accusations. This shift in framing the accusations suggests that celebrity news is not concerned with the truth, but what the celebrities hold for the truth. In doing so, “celebrity news reporters establish their credibility by way of demonstrating the quality of their access to the sources of celebrity news, rather than by their capacity to deliver verifiable, evidence-based reports” (Turner, 2014a, p. 149).

- (16) Last week, **Greta Gerwig**, who appeared in **Allen's 2012 film *To Rome With Love***, said: "If I had known then **what I know now**, I would not have acted in the film."

Telegraph, 20 January 2018

- (17) **Spike Lee** has **apologised after he defended Woody Allen over allegations of sexual abuse** made against him by Dylan Farrow.

Independent, 14 June 2020

- (18) The only false note and concession to **“cancel culture”** is his role with **Woody Allen**. **Timothée Chalamet**, in the race for the Oscars, **distances himself** from the director of *A Rainy Day in New York*, caught up in **accusations of incest by his adopted daughter**.²⁶

Figaro, 14 September 2021

²⁵ **De Scarlett Johansson à Pierce Brosnan** en passant par **Catherine Deneuve**, des **personnalités** ont pris la défense de Woody Allen, boycotté par les distributeurs aux États-Unis en raison d'**accusations d'abus sexuels**, à l'occasion de la projection de son nouveau film *Un jour de pluie à New York* (sortie en France le 18 septembre) au Festival de Deauville. « Woody Allen est au **Panthéon des réalisateurs**. Son travail fait sens », **a déclaré l'ex-007 Pierce Brosnan**.

²⁶ Seule **fausse note** et concession à la « **cancel culture** », son rôle chez **Woody Allen**. **Timothée Chalamet**, en pleine course aux Oscars, se **désolidariser**a du réalisateur d'*Un jour de pluie à New York*, rattrapé par les **accusations d'inceste de sa fille adoptive**.

3. Taboo language and celebrity

This section examines the framing of the incest taboo through the various linguistic avoidance strategies, i.e., taboo language. The incestuous dimension of the accusations is rarely openly stated. Journalists use the vague term *abuse*, euphemising the accusations of incest, as seen in examples (19) and (20). It is worth noting that the articles do not focus on the accusations but on Woody Allen's latest movies *Wonder Wheel* or *A Rainy Day in New York*, released in a post-#MeToo film industry. Accusations are mentioned in the background, summarised in one or half a sentence. It illustrates the mixing between facts and fiction, characteristic of celebrity news.

- (19) The **abuse allegation**, it is often forgotten, happened during his bitter split and child custody battle with Mia Farrow.

Independent, 9 March 2018

- (20) His setbacks date back to the resurgence of **suspicious of sexual abuse** in 1992 against his adoptive daughter Dylan Farrow when she was 7 years old.²⁷

Figaro, 11 June 2019

The term *abuse* can be considered a euphemistic umbrella term that covers various forms of assault. On a continuum of vagueness regarding euphemisms, the press uses stronger euphemisms, intensifying the blurriness in the description of the accusations, as in examples (21)-(24). These euphemisms reinforce the vagueness of the accusations by using terms such as *actions*, *allegations*, *claims*, *doings*, and *misconduct*, which can refer to any kind of deeds. In addition, example (24) shows a complex grammatical construction that obscures the kinship between Dylan Farrow and Woody Allen, contributing to the blurriness of the accusations and making it difficult to perceive their incestuous characteristics.

- (21) In 2014, Dylan published a damning op-ed in *The New York Times*, detailing her father's **inappropriate gestures** and denouncing the film industry's apathy.²⁸

Figaro, 28 November 2017

- (22) Allen continues to deny **the allegations**.

Guardian, 16 January 2018

²⁷ Ses déboires remontent à la résurgence des **soupçons d'abus sexuels** en 1992 contre sa fille adoptive Dylan Farrow lorsque celle-ci avait 7 ans.

²⁸ En 2014, Dylan publiera dans le *New York Times* une tribune à charge contre son père, relatant dans le détail les **gestes déplacés** de celui-ci et dénonçant l'apathie de l'industrie du cinéma.

- (23) Dylan's brother, Ronan, a reporter for *The New Yorker* and a prominent figure in the #MeToo revelations, which notably implicated producer Harvey Weinstein, published an article in 2017 denouncing Woody Allen's **doings** towards his sister.²⁹

Figaro, 11 June 2019

- (24) After the filming of Woody Allen's film "A Rainy Day in New York" and the **revelations of sexual misconduct** against the **director's adopted daughter by her father** [...].³⁰

Progrès, 20 June 2019

In addition, it is worth underlining that some articles from the French and British press previous to the release of the documentary "Allen vs. Farrow" in 2021, mentioning *Woody Allen* and *incest*, are the results of circumstances. Incest was indirectly mentioned in relation to Woody Allen. In other words, the incest taboo is not explicitly linked to Woody Allen. The first context concerns the eroticised portrayal of incest in films. In line with the previous findings in Chapter 5, examining the fictional depiction of incest, incest is here described as consensual and desired, and not as sexual abuse (Brey, 2022). Examples (25)-(27) further illustrate the ambiguity of the press in defining incest as a consensual yet transgressive relationship. In this context, Woody Allen is mentioned as a prominent figure of the Hollywood film industry.

- (25) It is the first time that a **Woody Allen's** film is watched without real pleasure. [...] An ounce of **incest** between the lines? There it is. Psychoanalysis is not forgotten, nor are the allusions to the classics of the time (the local cinema is screening *Winchester 73*).³¹

Figaro, 30 January 2018

- (26) We see it in the work and lives of **Woody Allen**, Philip Roth, or Picasso, [...] The same can't be said of his queasy and quasi-autobiographical *Murmur of the Heart* (1971), with its **unperturbed celebration of mother-son incest**.

Guardian, 26 May 2018

²⁹ Le frère de Dylan, Ronan, reporter pour *The New Yorker* et figure de proue des révélations #MeToo, mettant en cause notamment le producteur Harvey Weinstein, a publié en 2017 un article dénonçant les **agissements** de Woody Allen sur sa sœur.

³⁰ Après le tournage du film de Woody Allen *A Rainy day in New York* et les **révélations d'attouchements sexuels de la fille adoptive du réalisateur par son père** [...].

³¹ C'est la première fois qu'un Woody Allen se regarde sans vrai plaisir. [...] Une once d'**inceste** entre les lignes ? Et voilà. La psychanalyse n'est pas oubliée, ni les clin d'œil aux classiques de l'époque (le cinéma local projette *Winchester 73*).

- (27) Adored by his disciples, Martin Scorsese, **Woody Allen**, Wes Anderson, Richard Linklater and Alejandro Gonzalez Inarritu among them, [...] The actress here plays Karin, a schizophrenic recuperating with her family on a remote island, troubled by her emotionally distant father, **incestuous attraction to her brother** [...]

Independent, 27 July 2018

The second context shyly implies a connection between incest and the celebrity. The type of abuse is not explicitly said to be incestuous as it is characterised through vague terms such as *sexual assault* or *sexual misconduct* as in examples (28) and (29). The term *incest* is mentioned in parenthesis to explain the acronym of the American association RAINN (Rape, Abuse, and Incest National Network). In both examples, RAINN is mentioned because the Hollywood actor Timotée Chalamet who starred in Woody Allen's film *A Rainy Day in New York* donated his money to the association following the renewed accusations of sexual abuse by Dylan Farrow in 2021. It is worth noting that the prime focus is on the celebrity Chalamet, who is praised for his feminist gesture following #MeToo movement, while Dylan and her accusations are relegated into the background for context.

- (28) Colin Firth has joined the increasing number of actors who are turning their backs on **Woody Allen** following **allegations of sexual assault** made by his adopted daughter. [...] Earlier this week, Chalamet said: "I don't want to profit from my work on the film, and to that end I am going to donate my entire salary to three charities: Time's Up, the LGBT Centre in New York, and **[the Rape, Abuse & Incest National Network]**."

Telegraph, 20 January 2018

- (29) After the filming of **Woody Allen's** film *A Rainy Day in New York* and the **revelations of sexual misconduct** by the director's adopted daughter involving her father, Timothée Chalamet distanced himself from the actor-director and donated his earnings to three organizations: Time's Up, The LGBT Center in New York, and **RAINN (Rape, Abuse, and Incest National Network)**.³²

Progrès, 20 June 2019

Nevertheless, there are also explicit mentions of incest in relation to Woody Allen in the corpus. The avoidance linguistic strategies, circumvoluting the accusations against the celebrity, are primarily found before the release of the documentary "Allen vs. Farrow" in 2021. After this, the term *incest* is more frequently used to qualify the type of sexual abuse Woody Allen is accused of. As seen in examples (30)-(32), the term *incest* is no longer avoided

³² Après le tournage du film de Woody Allen *A Rainy day in New York* et les révélations d'attouchements sexuels de la fille adoptive du réalisateur par son père, Timothée Chalamet s'est désolidarisé de l'acteur-réalisateur et a reversé ses gains à trois associations : Time Up, The LGBT Center in New York et **RAINN (Rape, Abuse, and Incest National Network)**.

but explicitly mentioned. It suggests that the documentary has had a strong influence on how Dylan Farrow's accusations are perceived, since Dylan Farrow uses the term *incest* to describe herself as an *incest survivor* in the last documentary episode (Dick & Ziering, 2021). Thus, the journalists seem to reuse this term and acknowledge the seriousness of the accusations.

- (30) HBO was taking big risks by tackling, through a top-secret production, the subject of **incest** in the oldest, most obscure, and most controversial case in American cultural Olympus.³³

Libération, 26 February 2021

- (31) That a new, four-part documentary, which re-examines the allegations of **incest** - [...]. But she didn't, she says, ever think of herself as an **incest survivor**.

Times, 13 March 2021

- (32) The particular kind of abuse that Dylan alleges against Allen, **incest**, also has recognizable patterns. "**Incest**" is a term that often implies consent, or at least some sort of mutuality, but in reality, **incest** is what happens when the power imbalances inherent in the family structure are harnessed as a tool of sexual assault.

Guardian, 13 March 2021

In addition to the accusations of incest as child sexual abuse, Woody Allen is also known for marrying another adoptive daughter of his ex-partner, Mia Farrow. The newspapers mention his relationship with Soon-Yi, who he had known when she was a teenager, and emphasise the unconventional dimension of the relationship, as seen in examples (33)-(37). The terms *inappropriate*, *distasteful*, *unconventional*, *controversial*, *shocked*, and *scandal*, suggest the journalists' disapproval, or at least, judgement. Interestingly, this relationship, which seems to be consensual for both of them, is never characterised as incestuous, whereas it was noticed in Chapter 5 that consensual intrafamilial relationships are often the ones to be labelled as *incest*, especially in the British press. Thus, it is worth noting that the relationship with Soon-Yi, where the age difference is strongly emphasised, is framed as unconventional, but never as incestuous in this corpus. Interestingly, Dylan's accusations are sometimes linked to this relationship, like in examples (33) and (36). The relationship with Soon-Yi is tabooed, and the linguistic avoidance strategies are judgmental. Furthermore, hedges reinforce the linguistic avoidance strategies: "as distasteful as it was *at first*," "close to conventional," "loving if unconventional," "doesn't see anything *wrong*," "morally, it's a *different matter*."

- (33) Since 1992, Woody Allen's name has been associated with **suspicious of inappropriate touching** towards his adoptive daughter Dylan, who is now 32 years

³³ HBO prenait de gros risques en s'attaquant, au prix d'un tournage top secret, au sujet de l'**inceste** par le biais de l'affaire la plus ancienne, la plus opaque et la plus controversée de l'Olympe culturel américain.

old, **and the scandal of a relationship** with Soon-Yi Previn, the adoptive daughter of his ex-wife Mia Farrow... whom he married in 1997.³⁴

Figaro, 28 November 2017

- (34) No one in the entire documentary points out that Allen and Soon-Yi have now been together for almost 25 years, which suggests **their relationship - as distasteful as it was at first** - is clearly more than a smokescreen, and **now looks close to conventional**.

Guardian, 4 March 2021

- (35) According to Allen and his defenders, his art is fiction and not representative of his inner psyche, **his relationship with Soon-Yi Previn is loving if unconventional**.

Guardian, 13 March 2021

- (36) Back in the early 90s, people were **more shocked by his relationship with Soon-Yi than** they were by the soon scotched **allegation of child molestation**. [...] **explosive allegations of incest** and an esteemed auteur's **controversial relationship** with his partner's adopted daughter, 35 years his junior.

Times, 22 February 2021

- (37) It looks bad. **Woody Allen doesn't see anything wrong with sleeping with his partner's adopted daughter**. Legally, there's nothing to blame him for. **Morally**, it's a different matter.³⁵

Figaro, 28 February 2021

4. Summary

This case study suggests that celebrity news may cover sensational accusations because they become newsworthy, while avoiding their taboo nature through linguistic avoidance strategies. This underscores the paradoxical dimension of celebrity news: while the French and British press may contribute to breaking a widespread taboo by relaying this news, they also reinforce the taboo nature of the accusations through linguistic avoidance strategies.

In addition, what makes this type of news newsworthy is not the commonness of the taboo incest but that Woody Allen is accused of committing incestuous abuse. This observation

³⁴ Depuis 1992, le nom de Woody Allen est associé aux **souppçons d'attouchements** envers sa fille adoptive Dylan, aujourd'hui âgée de 32 ans, et au **scandale d'une liaison** avec Soon-yi Previn, la fille adoptive de son ex-femme Mia Farrow... qu'il a épousée en 1997.

³⁵ Ça fait désordre. **Woody Allen ne voit pas le mal qu'il y a à coucher avec la fille adoptive de sa compagne**. Légalement, on ne peut rien lui reprocher. **Moralement**, c'est une autre affaire.

should be nuanced as depictions of incestuous abuse that do not occur in “famous” families were observed in Chapters 5 and 6. However, it is worth noting that Woody Allen is the most frequently mentioned celebrity name in the French press, and even more frequently in the British press. Out of 50 articles in the English corpus, 80% were retrieved from broadsheets (*The Independent*, *The Guardian*, and *The Times*). To some extent, it may indicate the tabloidisation of the “quality” press. Nevertheless, despite the sensational dimension of such media coverage, celebrity news does impersonate a problem, which can help to further discuss the taboo. As such, sensationalism offers another perspective and is not necessarily the sign of the “dumbing down” of the press. The fact that Dylan Farrow is given a chance to express her voice almost thirty years after the first accusations, suggests that society is more open to listening. Indeed, the press coverage on Dylan Farrow’s accusations shows an interesting shift of perspective. From being passivated, she becomes the focus of the article. However, for this to happen, a documentary was released and emphasised her voice. It shows that celebrity news is a large system involving not only the press but documentaries and other media supports influencing one another. It demonstrates that newsworthiness evolves.

Nevertheless, this analysis reveals that both the accused and the accuser are more or less believed, not based on facts but on trendy perceptions. Celebrities’ opinions are either used to show their support, or to express their distancing. This mixing of facts and fiction contribute to blurring the seriousness of the accusations. Therefore, incest is not discussed *per se*, but because it involves a celebrity. By contrast, the fact that 1 family out of 6 in the United States is concerned with intrafamilial sexual abuse (Gilgun, 1995; Russell, 1983, 1999) is never mentioned. This further indicates that incestuous abuse in relation to a celebrity is mentioned because it is sensational, not because it may open up a larger social debate, demonstrating the extent of such sexual abuse, not limited to any social class. As such, the depiction of the celebrity tends to frame incest as an exceptional event, omitting to mention its ordinariness.

Chapter 8: Conclusion

Everything in the world is about sex, except sex. Sex is about power.

attributed to Oscar WILDE

This final chapter summarises the main findings on taboo language and incest in the French and British press from 2017 to 2022. The research question, consisting in determining the linguistic taboos on the incest taboo, is addressed by conceptualising the notion of taboo language outlined in this thesis. The implications of conducting cross-linguistic corpus-assisted discourse studies (CADS) are discussed, followed by potential future research that can be conducted to expand this thesis.

1. Linguistic taboos

The notion of taboo language is defined in this thesis as the linguistic avoidance strategies to discuss a taboo topic. In order to summarise the findings of this thesis, the notion of language taboo is conceptualised in Figure 8.1. Drawing on the analysis that euphemisms can be perceived on a progressive scale of blurriness (“floutage”) instead of being classified into semantic categories (Jamet, 2010), the notion of language taboo is conceptualised on a progressive scale of absence. In addition, this classification relies on Trinch’s (2001) analysis of Latino women reporting domestic abuse committed by their husbands to US police officers, highlighting how explicitly or implicitly the victims testify that marital rape occurred. Trinch (2001) categorises the interviews on a spectrum, ranging from “direct and explicit (legal, not euphemistic, and unambiguous)” to “inexplicit, nearly non-existent (possibly would not have been reported if interviewer had not asked)” (p. 582). While some women explicitly use the term *rape*, other women avoid the term and prefer phrases such as “being forced into sex,” calling into question whether they make this distinction for politeness reasons and cultural differences, or whether they genuinely consider that there are semantic differences between *rape* and “being forced into sex.” This research contrasts women’s descriptions of their domestic assaults on a progressive scale. Similarly, this thesis conceptualises taboo language on a continuum from explicitness to silence. The notion of explicit presence is the starting point of the continuum, gradually becoming blurred until it reaches vagueness and absence.

Figure 8.1 Conceptualisation of taboo language

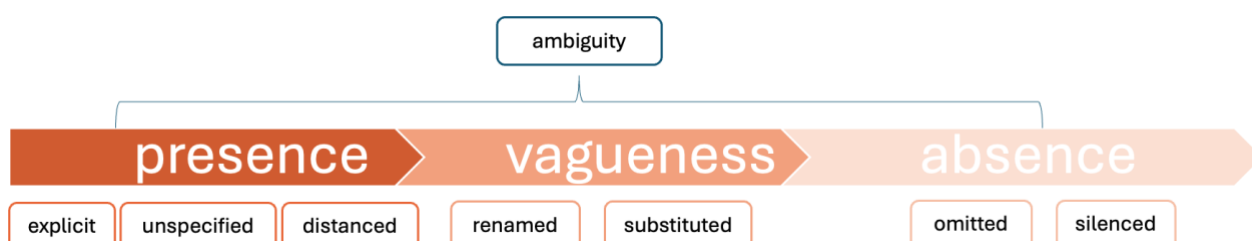


Figure 8. 1. relies on the findings of this thesis, but the conceptualisation of taboo language can be applied to other topics. This continuum may be simplified or expanded according to other researchers' findings. To further explain this notion of taboo language and summarise the findings of this research project, some examples illustrate each category on the continuum, going from unambiguous presence to complete absence.

Explicit presence refers to the clear and unambiguous mentions of a taboo topic. Chapter 5 suggests that the French press predominantly covers incest as child sexual abuse (CSA) with explicit terms referring to sexual violence such as *inceste* (incest), *père incestueux* (incestuous father), *violer* (rape), *violences sexuelles* (sexual violence), and *crime sexuel* (sexual crime). As such, incest is explicitly depicted as an intrafamilial crime against children, often underlining its large prevalence. Victims' voices are quoted to report their trauma and their need to testify. The analysis of the French corpus demonstrates that incest was framed as CSA over the five years under investigation, and even though the high-profile scandal in 2021 increased the newsworthiness of incest, the depictions of incest were predominantly explicit. In this category, there is no ambiguity: the taboo is explicitly discussed.

Unspecified presence relates to the unambiguous, yet brief and unexplained, mentions of a taboo topic. For instance, the term *incest* in the English corpus was mentioned amid a long enumeration of other terms in reference to the non-exception cases for accessing abortion in the US. Incest is explicitly mentioned but not discussed *per se*. The foregrounded topic of these articles is the abortion ban, not incest, as outlined in Chapter 5. Interestingly, the fact that the enumeration includes *rape* in more than half of the occurrences highlights the negative discourse prosody surrounding incest. On the other hand, it introduces some ambiguity. If *incest* is associated but distinguished from *rape*, what is implied through the term *incest*? Is the notion of rape implied, or is it excluded from the meaning of *incest*? Consequently, is consent implied or not implied in the term *incest*? These possible questions exemplify how incest is not explicitly defined. Similarly, in the Woody Allen corpus in Chapter 7, there were unspecified mentions of incest through the journalists' explanation of the acronym RAINN (standing for the American association Rape, Abuse and *Incest* National Network). While the article revolves around another celebrity distancing himself from Woody Allen, the incest nature of the accusations is sometimes never explicitly stated but is indirectly mentioned through the acronym. Therefore, both situations illustrate the unspecified presence of a taboo topic. The taboo is present but without being explicitly acknowledged or explained.

Distanced presence refers to the biased representation of a taboo, thus distancing its taboohood. Chapter 5 outlined how incest is predominantly covered through fictional representations in the English corpus, and in some rare occasions, in the French corpus. The eroticisation of incest, portrayed in fantasied worlds and pornography, tends to represent incest as a transgressive, yet consensual relationship. Its forbiddingness is arousing and becomes a plot element in fiction. However, this representation is ambiguous as it misrepresents, or even sublimates, the incest taboo. Thus, the depiction of the taboo

becomes ambivalent as its transgression is desired. In this sense, the taboo topic is present but distanced as it is misrepresented and ambiguous.

To some extent, the metaphorical mentions of incest may be classified as distanced presence. It was observed in Chapter 5 that the adjective *incestuous* is often used metaphorically to disapprove of promiscuity amid circles of power. This usage of *incestuous* can be considered a distanced presence since the notion of incest is mentioned but placed in a different context. Nevertheless, these findings pertain to a more abstract meaning of incest, applied to target domains that do not refer to sexuality. Thus, metaphors referring to taboo can be classified on this continuum or be considered separately.

From this explicit to distanced presence, the textual presence of a taboo topic becomes increasingly ambiguous, leading to vagueness. The first degree of vagueness can be labelled the **renamed presence**, i.e., a taboo topic is indirectly mentioned, so its presence is inferred from extratextual knowledge. For instance, in Chapter 6, the corpora “abuse AND father” and “abus AND père” show that incest could be renamed through other terms clearly stating sexual abuse, such as *rape*, *molest*, and *sexual abuse*, happening within the family circle, with terms like *father*, *grandfather*, *brother*, and *mother*. However, the term *incest* does not occur in the text. Therefore, the incestuous nature of abuse is indirectly mentioned, and can be inferred from context. Thus, the absence of clearly labelling this type of assault tends to make the description vague.

Substituted presence relates to the euphemistic terms used to discuss a taboo topic. For instance, Chapter 6 investigates how signalling nouns encapsulate phrases indirectly referring to incest through abstract notions. Terms like *this (kind of) abuse*, *these allegations*, *this claim*, and *this trauma*, do not only rename incest, but substitute its tabooess. This is why the substituted presence is considered one step further on the continuum compared to the renamed presence. These abstract terms indirectly refer to incest and contribute to the vagueness of the accusations reported in the press. The substituted presence euphemises the taboo topic and introduces ambiguity.

The taboo can shift from being vague to discursively absent. On this continuum of taboo language, absence comprises omitted and silenced presences. Both become noticeable through methodological comparisons as there is no textual presence of the taboo.

Omitted presence is a missing element that can be inferred from intertextual and contextual knowledge. In these findings, it corresponds to the knowledge that Woody Allen is accused of incestuous abuse by his adoptive daughter, as discussed in Chapter 7. Interestingly, some articles in the corpus did mention the incestuous nature of the abuse, whereas other articles were silent on this topic. Omitted presence can be noticed through the comparison of texts on the same topic, as some texts mention the taboo, while other texts within the same corpus do not. This omission may be intentional, done out of respect or politeness, such as avoiding the mention of someone’s trauma or protecting identities. In other cases, this absence can serve to hide the truth, such as preserving one’s reputation. Beyond these two possibilities –

one perceived as positive and the other as negative – uncertainty and/or lack of knowledge can be seen as an alternative explanation for omission. A journalist may avoid reporting a taboo topic because the information is uncertain and could be perceived as hearsay. It is difficult, or even impossible, to know the reasons why some texts may be willing or unwilling to report a taboo topic, as there is no metadata on the articles under investigation to explore who wrote them and why they chose to frame the topic in that manner.

Silenced presence refers to the complete absence of a taboo topic but is still noticeable through cross-textual comparisons. Thus, at the extreme end of this continuum, silenced presence refers to a more wide-scale absence compared to omitted presence. While omitted presence is observed through comparisons of texts within the same corpus, silenced presence is noticed through comparisons of corpora. For instance, the depiction of incest as child sexual abuse is absent in the English corpus but becomes noticeable when compared with the French corpus. Both corpora can be contrasted as they were built with the same lemma and during the same time period, as discussed in Chapter 5. Silenced presence becomes visible through data and/or method triangulations. As outlined in this thesis, the taboo is successful when it is completely absent. Thus, when a taboo topic is silenced, only analytical tools can make this absence visible. Furthermore, no ambiguity is introduced since there is no textual presence referring to any taboo whatsoever.

Thus, this taxonomy of linguistic taboos summarises the findings of this thesis in addressing the research question: What are the linguistic taboos on the tabooeness of incest? A subsidiary question was to investigate whether the linguistic taboos are similar or different when a taboo is socially shared between two cultures. As presented, the linguistic taboos on the tabooeness of incest are varied, and despite the universality of the incest taboo, the linguistic taboos tend to be culturally determined. The meaning of incest seems to be defined differently in France and the UK, mirrored in their respective press coverage. The term *incest* commonly refers to intrafamilial CSA in France, which is quite rarer in the UK. Interestingly, this finding is supported by other documents from the British authorities. For instance, a report investigating child sexual abuse within the family in the UK (Children’s Commissioner, 2015) mentions the term *incest* only once to be better dismissed in favour of the paraphrase *child sexual abuse in the family*: “research literature uses the phrase intra-familial child sexual abuse to refer to a variety of ‘familial/ incest’ dynamics and sexual behaviours, *broadly but not exactly equivalent* to the definition of *child sexual abuse in the family environment* adopted for this Inquiry” (p. 26, emphasis added). It is worth noting that no references are given to determine their “research literature,” obscuring the grounds on which they exclude the term *incest*. Furthermore, the distinction drawn between “*intrafamilial child sexual abuse*” and “*child sexual abuse in the family environment*” seems superficial as both terms are cognitively synonymous. However, another report on child sexual abuse in the UK (Beckett & Warrington, 2013) counts 491 offences in the category “*Incest or familial sexual offence*” (p. 38, emphasis added). Thus, the term *incest* can relate to (child) sexual abuse in the UK. However, this category is briefly shown in a table, thus, *incest* is also mentioned only once in this report. Therefore, it seems that incest can be understood as intrafamilial sexual abuse, but the rare

occurrences demonstrate that the British authorities partially (shyly?) acknowledge this definition. Without clear labelling, this social taboo might remain silenced for many more years.

2. Cross-linguistic considerations

This thesis aimed to contribute to cross-linguistic corpus-assisted discourse studies (CADS). The double triangulation between methods and data yielded significant results but was also challenged by unexpected technical aspects. One of the main practical challenges encountered in this cross-cultural thesis pertains to *anglonormativity*, i.e., “the naturalization and embeddedness of the English language” (Freake [Vessey], 2012, p. 5). Indeed, English is the predominant language in cross-linguistic CADS, as linguistic combinations often include English alongside another language: English/Chinese (Xiao & McEnery, 2014), English/German (Jaworska, 2021; Jaworska & Krishnamurthy, 2012), English/Italian (Taylor, 2014), and English/French (Baker & Vessey, 2022; Curry & Chambers, 2017). This thesis is no exception to this observation and encountered practical challenges due to English embeddedness in software. As outlined in Chapter 4, corpus linguistics is more developed for English than for French. A consequence of this discrepancy was visible through the automatic tagging of Sketch Engine, which made obvious mistakes in classifying words in French. For instance, the term *handicapé* (disabled) was considered a past participle and not an adjective, which caused Word Sketch to classify the verb *handicaper* (handicap) in the list of frequently occurring verbs, instead of classifying *handicapé* in the list of adjectives. This type of mistakes required manual removal.

Another consequence of *anglonormativity* regards the affordances of software that are only available in English. For instance, the interfaces of Lexis Nexis, Sketch Engine, Python, Deduplication Tool and Atlas.ti – all used in this thesis – are available only in English. This English embeddedness was not an obstacle to the research but raises practical issues such as outlined in Chapter 6, Section 1. Indeed, to build the corpora “abuse AND father” and “abus AND père,” the conjunction AND had to be written in English to instruct the software to retrieve articles containing both terms.

These practical details were addressed, for instance by using the English coding on Lexis Nexis for the French search query or excluding incorrectly tagged French data on Sketch Engine. However, this suggests that English embeddedness facilitates research in English and focuses on English data and findings. Therefore, more sophisticated tools are needed for conducting multilingual research. Nevertheless, beyond these considerations, it should be acknowledged that mastering software such as Lexis Nexis, Python and Atlas.ti, used for collecting, sorting and ordering data, often required assistance from other researchers and/or instructors. Regardless of its language, handling a large amount of data in corpus linguistics requires good IT skills, which can be time-consuming to master (Brezina, 2018).

An additional practical challenge in cross-linguistic CADS relates to keyword analysis. As outlined in Chapters 4 and 5, comparing keywords between two corpora of different languages is close to impossible because the reference corpus is necessarily different, providing different results. For this reason, the analysis focused on the most frequent lexical nouns, which also gave an understanding of the topicality of each corpus, as demonstrated through the prevalence of the topic of abortion in the English corpus, compared to the depiction of incest as child sexual abuse in the French corpus. Thus, this thesis used keyword lists to corroborate the lists of frequent lexical nouns. As such, frequency analyses seem to be a more robust and reliable approach in cross-linguistic CADS. Instead of differentiating corpora through keywords, and thus highlighting differences over similarities, frequency lists favour results generated by each corpus.

3. Future directions

This study could be further expanded by investigating the visual representations of incest in the press. The pictures of the articles regarding incest were not included in this research that chose to focus on the linguistic patterns of the incest taboo. Nevertheless, it was noticed that many French articles were accompanied by a picture of a child, afraid and alone. As Schwark (2017) points out regarding the visual representations of sexual violence against women, the reader observes the scene from the aggressor's point of view. For instance, pictures illustrating articles reporting rape show women alone in dark streets where the reader is often in the position of the aggressor, lurking behind. Similarly, the pictures illustrating articles on incest place the reader as the aggressor: in front of children, afraid and alone, and shielding their face with their hands. They are portrayed in dark and unrecognisable places. These visual depictions are somehow far from reality as incestuous abuse happens in familiar environments (house, bedroom, bathroom). Furthermore, representing children in this way contributes to depicting them as defenceless, which is partly accurate but also stereotyped. As outlined in Chapter 2, children are socially construed as defenceless, thus granting adults more rights than them (Piterbraut-Merx, 2022).

One step further into visual representations would be to investigate intelligence-generated images to perceive how incest is represented based on a large amount of visual data. Research demonstrates that AI text-to-image generation reproduces social biases and stereotypes (Bianchi et al., 2023; Putland et al., 2023). It is worth exploring whether incest would be predominantly represented as child sexual abuse or eroticised relationship. Furthermore, if possible, it would be interesting to contrast images from French-speaking sources and English-speaking sources. Drawing on the above observations on anglonormativity, the analysis could address whether a dominant linguistic dataset would create a predominant visual representation, obfuscating other possible representations generated from different linguistic areas. Based on the findings of this thesis, it would be worth observing whether the English depiction of incest, emphasising erotic/transgressive consent between two adults, would be

the main representation, or whether the French perception of incest as child sexual abuse, would be present or absent.

Future research could incorporate forensic linguistic analyses, focusing on the defence and prosecution discourses in incest trials. For instance, Rosulek (2008) investigates the linguistic manipulations in the closing remarks, observing what is foregrounded or backgrounded. Her research focuses on child sexual abuse but does not address incest specifically, which could be examined separately. In line with research conducted by Gilgun (1995) and Dussy (2021), another angle could examine the incest perpetrators' discourses to understand how they portray themselves and their victims. Conversely, analysing victims' discourses can highlight their points of view. For instance, victims took the floor at public meetings organised by the Committee of Inquiry on incest and sexual violence against children in France from 2021 to 2023. During these meetings, victims spoke about their lasting trauma and struggles. Sometimes, their speeches were spontaneous, sometimes, they were written. Some speeches pointed out the healing process through the power of speaking. At a public meeting in Paris, attended by this author on 16 February 2022, a victim said: "With incest, we are in the unthinkable, the unspeakable, but still, we have to find words", and another one: "By coming here, it's like a therapy. As soon as I speak, it feels good." These testimonies highlight the victims' consciousness of breaking the taboo through language. Such linguistic analyses could give voice to them.

4. Concluding remarks

I would like to finish this thesis on a more personal note. The idea of this thesis came from the bewilderment of discovering the large prevalence of incest in French society following the publication of the book *La Familia grande* (Kouchner, 2021). Thus, beyond the theoretical approaches of finding discursive absence, this thesis aims to shed light on incestuous abuse that has been investigated for over 40 years, predominantly by scholars from the United States and Canada in the 1980s, as mentioned in Chapter 2. Despite years of research in anthropology, sociology, and psychology, the incest taboo is still silenced and denied.

In France, the Committee of Inquiry on incest and sexual violence against children estimated that the financial consequence of not tackling child sexual abuse costs around 9.7 billion euros, largely due to the health system costs for victims (CIIVISE, 2023). The society's choice of tackling the prevalence of sexual abuse or of denying its existence reveals how children are considered (Porter & Ciba Foundation, 1984). The judge Édouard Durand, who had the co-charge of the Committee of Inquiry for over two years, emphasises the impossibility for victims to speak up. When they talk, they are said to lie, but when they do not speak, they are blamed for staying silent (Durand, 2024). The strength of denial, from the family and society at large, allows aggressors to feel powerful, and thus, incestuous abuse is perpetuated. Therefore, investigating incest poses the question of who holds power and who is protected

through the law. It is worth highlighting that only 3% of pedocriminals³⁶ are found guilty in France. Durand (2024) notes: “Deciphering the mechanisms of sexual violence and its extent, the mechanisms of denial and the perpetrators’ impunity, is to question the very notion of power and the legitimacy of those who have seized it” (p. 24).

Notably, this judge, acclaimed for his relevant and efficient work, has been dismissed from his presiding duties in December 2023 and the Committee is now progressively shushed down (Ouest-France, 2024). It suggests that there is no strong political will to tackle sexual violence within the family, both in France and the UK. It poses the question whether the representations of loving and caring families are aimed to be preserved for our societies’ sake. Indeed, I consider, and progressively so along this research, that exploring incest questions the basis on which French and British societies (and many more) rely: the nucleus family, often presented as the guarantee of happiness. Thus, the prevailing silence on the incest taboo reinforces social misconceptions. I hope that this thesis, on its modest scale, contributes to breaking the incest taboo.

³⁶ In light of the findings, this literal translation into English from the French term *pédocriminel* is selected over *paedophile*.

References

À propos d'un procès. (1977). *Le Monde*.

https://www.lemonde.fr/archives/article/1977/01/26/a-propos-d-un-proces_2854399_1819218.html

Alcántara-Plá, M., & Ruiz-Sánchez, A. (2018). Not for Twitter: Migration as a Silenced Topic in the 2015 Spanish General Election. In M. Schröter & C. Taylor (Eds.), *Exploring Silence and Absence in Discourse, Empirical Approaches* (pp 25-64). Palgrave.

Algava, É., Bloch, K., & Robert-Bobée, I. (2021). *Les familles en 2020: 25 % de familles monoparentales, 21 % de familles nombreuses*. Institut national de la statistique et des études économiques.

Allan, K. (Ed.). (2019). *The Oxford handbook of taboo words and language*. Oxford University Press.

Allan, K., & Burrige, K. (1991). *Euphemism & dysphemism: Language used as shield and weapon*. Oxford University Press.

Allan, K., & Burrige, K. (2006). *Forbidden words: Taboo and the censoring of language*. Cambridge University Press.

Allen, W., Kapsis, R. E., & Coblentz, K. (2006). *Woody Allen: Interviews*. University Press of Mississippi.

Ambroise-Rendu, A.-C. (2016). Briser le tabou: Du secret à la parole médiatique, le tournant des années 1970-1990. *Sociétés & Représentations*, 42(2), 59.
<https://doi.org/10.3917/sr.042.0059>

- Ancarno, C. (2018). Interdisciplinary approaches in corpus linguistics and CADS. In C. Taylor & A. Marchi (Eds.), *Corpus approaches to discourse: A critical review*. Routledge.
- Anderson, B. R. O. (2016). *Imagined communities: Reflections on the origin and spread of nationalism* (Revised edition). Verso.
- Angot, C. (2013). *L'inceste*. J'ai lu.
- Angot, C. (2017). *Un amour impossible*. Éditions Gallimard.
- Armstrong, L. (1978). *Kiss Daddy goodnight: A speak-out on incest*. Hawthorn Books.
- Asr, F. T., Mazraeh, M., Lopes, A., Gautam, V., Gonzales, J., Rao, P., & Taboada, M. (2021). The Gender Gap Tracker: Using Natural Language Processing to measure gender bias in media. *PLOS ONE*, 16(1), e0245533. <https://doi.org/10.1371/journal.pone.0245533>
- Austin, J. L. (1991). *Quand dire, c'est faire* (G. Lane, Trans.). Éd. du Seuil.
- Awad, G. (1995). *Du sensationnel : Place de l'événementiel dans le journalisme de masse*. Harmattan.
- Baker, P. (2004). Querying Keywords: Questions of Difference, Frequency, and Sense in Keywords Analysis. *Journal of English Linguistics*, 32(4), 346–359. <https://doi.org/10.1177/0075424204269894>
- Baker, P. (2006). *Using corpora in discourse analysis*. Continuum.
- Baker, P. (2010). *Sociolinguistics and corpus linguistics*. Edinburgh University Press.
- Baker, P. (2023). *Using corpora in discourse analysis* (Second edition). Bloomsbury Academic.

- Baker, P., & Egbert, J. (Eds.). (2016). *Triangulating methodological approaches in corpus-linguistic research*. Routledge.
- Baker, P., Gabrielatos, C., KhosraviNik, M., Krzyżanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society*, 19(3), 273–306. <https://doi.org/10.1177/0957926508088962>
- Baker, P., Gabrielatos, C., & McEnery, T. (2013). *Discourse analysis and media attitudes: The representation of Islam in the British press*. Cambridge University Press.
- Baker, P., & Levon, E. (2015). Picking the right cherries? A comparison of corpus-based and qualitative analyses of news articles about masculinity. *Discourse & Communication*, 9(2), 221–236. <https://doi.org/10.1177/1750481314568542>
- Baker, P., & Vessey, R. (2022). A corpus-driven comparison of English and French Islamist extremist texts. *International Journal of Corpus Linguistics*, 255–278. <https://doi.org/10.1075/ijcl.17108.bak>
- Baker, P., Vessey, R., & McEnery, T. (2021). *The language of violent jihad*. Cambridge University Press.
- Barnett, A. (2020). A genealogy of hostility: Parental alienation in England and Wales. *Journal of Social Welfare and Family Law*, 42(1), 18–29. <https://doi.org/10.1080/09649069.2019.1701921>
- Béart, E., & Mikova, A. (Directors). (2023). *Un silence si bruyant* [Broadcast]. M6.
- Beckerman, S., & Valentine, P. (Eds.). (2002). *Cultures of multiple fathers: The theory and practice of partible paternity in lowland South America*. Univ. Press of Florida.

- Beckett, H., & Warrington, C. (2013). *Suffering in silence: Children and unreported crime*. All Party Parliamentary Group for Victims and Witnesses of Crime.
- Bednarek, M. (2019). 'Don't say crap. Don't use swear words.' – Negotiating the use of swear/taboo words in the narrative mass media. *Discourse, Context & Media*, 29, 100293. <https://doi.org/10.1016/j.dcm.2019.02.002>
- Bednarek, M., & Caple, H. (2014). Why do news values matter? Towards a new methodological framework for analysing news discourse in Critical Discourse Analysis and beyond. *Discourse & Society*, 25(2), 135–158. <https://doi.org/10.1177/0957926513516041>
- Bednarek, M., & Caple, H. (2017). *The discourse of news values: How news organizations create 'newsworthiness'*. Oxford University Press.
- Bednarek, M., & Caple, H. (2019). *News discourse*. Bloomsbury Academic.
- Bennett, C., & Harden, J. (2019). Sexuality as taboo: Using interpretative phenomenological analysis and a Foucauldian lens to explore fathers' practices in talking to their children about puberty, relationships and reproduction. *Journal of Research in Nursing*, 24(1–2), 22–33. <https://doi.org/10.1177/1744987118818863>
- Bianchi, F., Kalluri, P., Durmus, E., Ladhak, F., Cheng, M., Nozza, D., Hashimoto, T., Jurafsky, D., Zou, J., & Caliskan, A. (2023). Easily Accessible Text-to-Image Generation Amplifies Demographic Stereotypes at Large Scale. *2023 ACM Conference on Fairness, Accountability, and Transparency*, 1493–1504. <https://doi.org/10.1145/3593013.3594095>
- Billon, A., Borchio Fontimp, A., Cohen, L., & Rossignol, L. (2022). *Porno: L'enfer du décor* [Rapport d'information]. Sénat.

- Biressi, A., & Nunn, H. (Eds.). (2008). *The tabloid culture reader*. Open University Press.
- Borneman, J. (2001). Caring and Being Cared For: Displacing Marriage, Kinship, Gender, and Sexuality. In J. Faubion, *The Ethics of Kinship: Ethnographic Inquiries*. Rowman & Littlefield Publishers.
- Bourdieu, P. (1996). *Sur la télévision: Suivi de L'emprise du journalisme*. Liber éditions.
- Bourdieu, P. (2001). *Langage et pouvoir symbolique*. Ed. du Seuil.
- Bourdin, D. (2005). 'L'interprétation des rêves', Sigmund Freud avec le texte du chapitre VI [paragraphe] 1, 2 et 3. Bréal.
- Brandes, S. H. (2018). An anthropological approach to taboo words and language. In K. Allan (Ed.), *The Oxford Handbook of Taboo Words and Language* (pp. 371–390). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198808190.013.20>
- Brey, I. (2022). L'inceste qui crève les yeux. Représentations dans les séries et au cinéma. In *La Culture de l'inceste*. Éditions du Seuil.
- Brey, I., & Drouar, J. (2022). *La culture de l'inceste*. Éditions du Seuil.
- Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide*. Cambridge University Press.
- Broder, A. Z. (2000). Identifying and Filtering Near-Duplicate Documents. In R. Giancarlo & D. Sankoff (Eds.), *Combinatorial Pattern Matching* (Vol. 1848, pp. 1–10). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-45123-4_1
- Brookes, G., & Baker, P. (2021). *Obesity in the news: Language and representation in the press*. Cambridge University Press.

- Brown, D. K., Harlow, S., García-Perdomo, V., & Salaverría, R. (2018). A new sensation? An international exploration of sensationalism and social media recommendations in online news publications. *Journalism*, 19(11), 1497–1516. <https://doi.org/10.1177/1464884916683549>
- Brown, K. (2023). *Talking 'with' and 'about' the far right: Putting the mainstream in mainstreaming*. PhD Thesis.
- Brownmiller, S. (1993). *Against our will: Men, women, and rape*. Fawcett Columbine.
- Cadoret, A. (2007). L'homoparenté: Un révélateur de l'ordre familial ? *Recherches familiales*, 4(1), 47. <https://doi.org/10.3917/rf.004.0047>
- Cameron, D. (1995). *Verbal hygiene*. Routledge.
- Cameron, D. (2001). *Feminism and linguistic theory*. Palgrave.
- Cantarella, E., & Sirignano, P. (2019). *Les plus belles histoires d'amour de l'Antiquité: Du ciel à la Terre, de Zeus à César*. Albin Michel.
- Caple, H., Anthony, L., & Bednarek, M. (2019). *Kaleidographic: A data visualization tool*. *International Journal of Corpus Linguistics*, 24(2), 245–261. <https://doi.org/10.1075/ijcl.18038.cap>
- Carrive, L. (2021). #MeTooInceste: Près de 80 000 messages en deux jours. *France Inter*. <https://www.radiofrance.fr/franceinter/metooinceste-pres-de-80-000-messages-en-deux-jours-5899685>

Carter, C. (1998). When the 'extraordinary' becomes 'ordinary', Everyday news of sexual violence. In C. Carter, G. Branston, & S. Allan (Eds.), *News, gender, and power*. Routledge.

Carter, C., Branston, G., & Allan, S. (Eds.). (1998). *News, gender, and power*. Routledge.

Chamizo-Domínguez, P. J. (2018). Ambiguity and vagueness as cognitive tools for euphemistic and politically correct speech. In A. P. Pedraza, *Linguistic Taboo Revisited: Novel Insights from Cognitive Perspectives*. De Gruyter.

Channell, J. (1994). *Vague language*. Oxford University Press.

Charles, M. (2003). 'This mystery...': A corpus-based study of the use of nouns to construct stance in theses from two contrasting disciplines. *Journal of English for Academic Purposes*, 2(4), 313–326. [https://doi.org/10.1016/S1475-1585\(03\)00048-1](https://doi.org/10.1016/S1475-1585(03)00048-1)

Children's Commissioner. (2015). *Protecting children from harm: A critical assessment of child sexual abuse in the family network in England and priorities for action*.

Chilton, P. (2005). Missing links in mainstream CDA: Modules, blends and the critical instinct. In R. Wodak & P. Chilton (Eds.), *Discourse Approaches to Politics, Society and Culture* (Vol. 13, pp. 19–51). John Benjamins Publishing Company. <https://doi.org/10.1075/dapsac.13.05chi>

CIIVISE. (2022). *Violences sexuelles: Protéger les enfants* (Conclusions intermédiaires).

CIIVISE. (2023). *Violences sexuelles faites aux enfants: Le coût du déni*. Commission indépendante sur l'inceste et les violences sexuelles faites aux enfants.

Cocteau, J. (Director). (1963). *La Machine infernale* [Video recording].

- Colman, A. M. (2015). Genetic sexual attraction. In *A dictionary of psychology* (Fourth edition). Oxford University Press.
- Cotterill, J. (2007). 'I Think He Was Kind of Shouting or Something': Uses and Abuses of Vagueness in the British Courtroom. In J. Cutting, *Vague Language Explored*. Palgrave Macmillan.
- Couldry, N., & Markham, T. (2007). Celebrity culture and public connection: Bridge or chasm? *International Journal of Cultural Studies*, 10(4), 403–421. <https://doi.org/10.1177/1367877907083077>
- Crespo-Fernández, E. (2015). *Sex in Language: Euphemistic and Dysphemistic Metaphors in Internet Forums*. Bloomsbury Academic. <https://doi.org/10.5040/9781474226851>
- Crespo-Fernández, E. (2018). Taboos in speaking of sex and sexuality. In K. Allan & K. Burridge, *The Oxford Handbook of Taboo Words and Language*. Oxford University Press.
- Cumbria County Council. (n.d.). *Genetic sexual attraction*.
- Curry, N., & Chambers, A. (2017). Questions in English and French Research Articles in Linguistics: A Corpus-Based Contrastive Analysis. *Corpus Pragmatics*, 1(4), 327–350. <https://doi.org/10.1007/s41701-017-0012-0>
- Cutting, J. (Ed.). (2007). *Vague language explored*. Palgrave Macmillian.
- De Benedictis, S., Orgad, S., & Rottenberg, C. (2019). #MeToo, popular feminism and the news: A content analysis of UK newspaper coverage. *European Journal of Cultural Studies*, 22(5–6), 718–738. <https://doi.org/10.1177/1367549419856831>

de Foucher, L. (2021). « La France a un problème avec l'inceste » : Avec #metooinceste, des dizaines de milliers de Tweet libèrent la parole. *Le Monde*.
https://www.lemonde.fr/societe/article/2021/01/18/la-france-a-un-probleme-avec-l-inceste-avec-metooinceste-des-dizaines-de-milliers-de-tweets-liberent-la-parole_6066605_3224.html

Dick, K., & Ziering, A. (Directors). (2021). *Allen vs. Farrow* [Broadcast]. HBO.

Diffloth, G. (2014). To Taboo Everything at All Times. *Annual Meeting of the Berkeley Linguistics Society*, 6. <https://doi.org/10.3765/bls.v6i0.2141>

Dijk, T. A. van. (1988). *News as discourse*. L. Erlbaum associates.

Douglas, M. (2005). *Purity and danger: An analysis of concept of pollution and taboo*. Routledge.

Dubied, A., & Hanitzsch, T. (2014). Studying celebrity news. *Journalism*, 15(2), 137–143.
<https://doi.org/10.1177/1464884913488717>

Duguid, A., & Partington, A. (2018). Absence, You don't know what you're missing. Or do you? In C. Taylor & A. Marchi (Eds.), *Corpus Approaches to Discourse, A Critical Review*. Routledge.

Durand. (2024). *160 000 enfants : Violences sexuelles et déni social*. Gallimard.

Durkheim, É. (1987). *La prohibition de l'inceste et ses origines: Étude de sociologie*. Éditions Payot & Rivages.

Dussy, D. (2022). Du rôle des « grands hommes » dans la reconduction des pratiques de l'inceste. In I. Brey & J. Drouar, *La Culture de l'inceste*. Editions du Seuil.

- Dussy, D. (with Pudlowski, C.). (2021). *Le berceau des dominations: Anthropologie de l'inceste* (Revised Edition). Pocket.
- Dyoniziak, J. (2018). L'euphémisme face à la construction de l'identité sociale. Le cas du discours médiatique. *Acta Universitatis Lodziensis. Folia Litteraria Romanica*, 12, 97–111. <https://doi.org/10.18778/1505-9065.12.10>
- Ehrlich, S., Meyerhoff, M., & Holmes, J. (Eds.). (2014). *The handbook of language, gender, and sexuality* (Second edition). Wiley-Blackwell.
- Elliott, M. (Ed.). (1993). *Female sexual abuse of children: The ultimate taboo*. Longman.
- Enslar, E. (2004). *The Vagina Monologues*. Virago Press.
- Fairclough, N. (2013). *Critical discourse analysis: The critical study of language*. Routledge.
- Farrow, D. (2014). An Open Letter From Dylan Farrow. *The New York Times*. <https://archive.nytimes.com/kristof.blogs.nytimes.com/2014/02/01/an-open-letter-from-dylan-farrow/?mcubz=1>
- Farrow, D. (2017). Why has the #MeToo revolution spared Woody Allen? *Los Angeles Times*. <https://www.latimes.com/opinion/op-ed/la-oe-farrow-woody-allen-me-too-20171207-story.html>
- Farrow, R. (2017). From Aggressive Overtures to Sexual Assault: Harvey Weinstein's Accusers Tell Their Stories. *The New Yorker*. <https://www.newyorker.com/news/news-desk/from-aggressive-overtures-to-sexual-assault-harvey-weinsteins-accusers-tell-their-stories>

- Ferrarotti, L. (2011). The news presenter and the television audience: A comparative perspective of the use of we and you. In L. Haarman & L. Lombardo, *Evaluation and Stance in War News*. Continuum.
- Flowerdew, J. (2002). *A Pedagogic Grammar of Signalling Nouns in Discourse*.
- Flowerdew, J., & Forest, R. W. (2015). *Signalling nouns in English: A corpus-based discourse approach*. Cambridge University Press.
- Foubert, O., Marinato, L., & Vallery, R. (2022). *Liberation of speech or liberation of listening? #MeTooInceste testimonies and their reception on Twitter*. Language, Ideology and Power research group seminar, Lancaster University.
- Fowler, R. (1991). *Language in the news: Discourse and ideology in the press*. Routledge.
- Francis, G. (2002). Labelling discourse: An aspect of nominal-group lexical cohesion. In M. Coulthard, *Advances in Written Text Analysis*. Taylor and Francis.
- Franklin, B. (1997). *Newszak and news media*. Arnold.
- Freake, R. (2012). A cross-linguistic corpus-assisted discourse study of language ideologies in Canadian newspapers. In *Proceedings of the 2011 Corpus Linguistics Conference*. Birmingham University: 1-11.
- Freake, R., Gentil, G., & Sheyholislami, J. (2011). A bilingual corpus-assisted discourse study of the construction of nationhood and belonging in Quebec. *Discourse & Society*, 22(1), 21–47. <https://doi.org/10.1177/0957926510382842>
- Freud, S. (1913). *Totem et tabou* (C. Leguil, Ed.; D. Tassel, Trans.). Seuil.
- Freud, S. (2013). *L'interprétation du rêve* (J.-P. Lefebvre, Ed.). Éd. du Seuil.

- Friginal, E., & Hardy, J. A. (Eds.). (2021). *The Routledge handbook of corpus approaches to discourse analysis*. Routledge.
- Gabrielatos, C. (2007). Selecting query terms to build a specialised corpus from a restricted-access database. *ICAME Journal*, 31, 5–43.
- Galli de' Paratesi, N. (2009). Eufemismo e disfemismo nel linguaggio politico e nell'italiano di oggi. *Synergies Italie*, 137–144.
- Game of Thrones*. (2011, 2019). [Video recording]. HBO.
- Garau, F. (2023). 'È pornografia'. Prof licenziata per la lezione sul David di Michelangelo. // *Giornale.It*. <https://www.ilgiornale.it/news/cronaca-internazionale/david-michelangelo-diventa-pornografico-docente-licenziata-2130688.html>
- Gardner, R. A. (1987). The parental alienation syndrome and the differentiation between fabricated and genuine child sex abuse. *Cresskill, NJ: Creative Therapeutics*.
- Gerding Speno, A., & Aubrey, J. S. (2018). Sexualization, Youthification, and Adultification: A Content Analysis of Images of Girls and Women in Popular Magazines. *Journalism & Mass Communication Quarterly*, 95(3), 625–646.
- <https://doi.org/10.1177/1077699017728918>
- Gilgun, J. F. (1995). We Shared Something Special: The Moral Discourse of Incest Perpetrators. *Journal of Marriage and the Family*, 57(2), 265. <https://doi.org/10.2307/353682>
- Gill, R., & Orgad, S. (2018). The shifting terrain of sex and power: From the 'sexualization of culture' to # MeToo. *Sexualities*, 21(8), 1313–1324.
- <https://doi.org/10.1177/1363460718794647>

- Gillings, M., Mautner, G., & Baker, P. (2023). *Corpus-Assisted Discourse Studies* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781009168144>
- Gómez, M. C. (2018). Lexicon, discourse and cognition: Terminological delimitations in the conceptualizations of linguistic taboo. In A. P. Pedraza, *Linguistic Taboo Revisited: Novel Insights from Cognitive Perspectives*. De Gruyter, Inc.
- Goudaillier, J.-P. (2018). Contournement des tabous: Exemples d'euphémismes et de dysphémismes dans les désignations en français de maladies graves et de la mort. *Acta Universitatis Lodzianis. Folia Litteraria Romanica*, 12, 71–80. <https://doi.org/10.18778/1505-9065.12.07>
- Government Statistical Service. (2011). *National Statistician's Review of Crime Statistics: England and Wales*. Government Statistical Service.
- Grabowski, Ł. (2022). Provoke or encourage improvements? On semantic prosody in English-to-Polish translation. *Perspectives*, 30(1), 120–136. <https://doi.org/10.1080/0907676X.2021.1891268>
- Greer, C. (2003). *Sex crime and the media: Sex offending and the press in a divided society*. Willan Pub.
- Gripsrud, J. (2000). Tabloidization, Popular Journalism, and Democracy. In C. Sparks & J. Tulloch (Eds.), *Tabloid tales: Global debates over media standards*. Rowman & Littlefield Publishers.
- Grogan, C. (2016). *Father-daughter incest in twentieth-century American literature: The complex trauma of the wound and the voiceless*. Fairleigh Dickinson University Press ; Rowman & Littlefield Publishing Group.

- Gruyer, F., Fadier-Nisse, M., & Sabourin, P. (1991). *La violence impensable: Inceste et maltraitance*. Nathan.
- Guibal, C. (2010). Toutankhamon, l'inceste fait pharaon. *Libération*.
https://www.liberation.fr/culture/2010/02/18/toutankhamon-l-inceste-fait-pharaon_610667/
- Guindi, F. E., & Read, D. (2012). Westermarck Hypothesis Reconsidered: A Comment on Kushnick and Fessler. *Current Anthropology*, 53(1), 134–135.
<https://doi.org/10.1086/663576>
- Gumbel, A. (2003). Dr Richard A. Gardner, Child psychiatrist who developed the theory of Parental Alienation Syndrome. *Independent*.
<https://www.independent.co.uk/news/obituaries/dr-richard-a-gardner-36582.html>
- Haarman, L., & Lombardo, L. (2011). *Evaluation and stance in war news: A linguistic analysis of American, British and Italian television news reporting of the 2003 Iraqi war*. Continuum.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Taylor and Francis.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *An introduction to functional grammar* (3rd ed). Oxford University Press.
- Harcup, T., & O'Neill, D. (2001). What Is News? Galtung and Ruge revisited. *Journalism Studies*, 2(2), 261–280. <https://doi.org/10.1080/14616700118449>
- Hardt-Mautner, G. (1995). *'Only Connect.'* *Critical Discourse Analysis and Corpus Linguistics*.

- Harkins, G. (2009). *Everybody's family romance: Reading incest in neoliberal America*. University of Minnesota Press.
- Héritier, F. (2012). *Les deux sœurs et leur mère : Anthropologie de l'inceste*. O. Jacob.
- Hetherington, J. (1999). The idealization of women: Its role in the minimization of child sexual abuse by females. *Child Abuse & Neglect*, 23(2), 161–174. [https://doi.org/10.1016/S0145-2134\(98\)00119-7](https://doi.org/10.1016/S0145-2134(98)00119-7)
- Hirsch, D. (2020). Les interdits, discriminants du corps et des affects à l'adolescence. *Le Carnet PSY*, N° 233-234(3), 25–31. <https://doi.org/10.3917/lcp.233.0025>
- Ho, C. (2024). *Saying the Unsayable*.
- Hoey, M. (2005). *Lexical priming: A new theory of words and language*. Routledge.
- Hrdy, S. B. (2011). *Mothers and Others: The Evolutionary Origins of Mutual Understanding*. Harvard University Press.
- Hunston, S. (2006). *Corpora in applied linguistics*. Cambridge University Press.
- Hunston, S. (2011). *Corpus approaches to evaluation: Phraseology and evaluative language*. Routledge.
- Hunston, S., & Francis, G. (2000). *Pattern grammar: A corpus-driven approach to the lexical grammar of English*. Benjamins.
- Hunter, R., Burton, M., & Trinder, L. (2020). *Assessing Risk of Harm to Children and Parents in Private Law Children Cases*. Ministry of Justice.

- Ivanič, R. (1991). Nouns in search of a context: A study of nouns with both open- and closed-system characteristics. *IRAL - International Review of Applied Linguistics in Language Teaching*, 29(2). <https://doi.org/10.1515/iral.1991.29.2.93>
- Jackson, C., & Tinkler, P. (2007). 'Ladettes' and 'Modern Girls': 'Troublesome' Young Femininities. *The Sociological Review*, 55(2), 251–272. <https://doi.org/10.1111/j.1467-954X.2007.00704.x>
- Jakubíček, M., Kilgarriff, A., Kovář, V., Rychlý, P., & Suchomel, V. (2013). The TenTen Corpus Family. In A. Hardie & R. Love (Eds.), *Corpus Linguistics 2013, Abstract Book*.
- Jamet, D. (2010). *Historique et procédés linguistiques de l'euphémisme*.
- Jamet, D., & Jobert, M. (with Centre d'études linguistiques). (2010). *Empreintes de l'euphémisme: Tours et détours*. l'Harmattan.
- Jaworska, S. (2014). The quest for the 'local' and 'authentic': Corpus-based explorations into the discursive constructions of tourist destinations in British and German commercial travel advertising. In D. Höhmann, *Tourismuskommunikation*. Peter Lang. <https://doi.org/10.3726/978-3-653-04165-1>
- Jaworska, S. (2021). Investigating media representations of the coronavirus in the UK, USA and Germany: What can a comparative corpus-based discourse analysis contribute to our understanding of the Covid-19 pandemic? In R. H. Jones, *Viral discourse*. Cambridge University Press.
- Jaworska, S., & Krishnamurthy, R. (2012). On the F word: A corpus-based analysis of the media representation of feminism in British and German press discourse, 1990–2009. *Discourse & Society*, 23(4), 401–431. <https://doi.org/10.1177/0957926512441113>

- Jaworska, S., & Leuschner, T. (2018). Crossing languages – crossing discourses: A corpus-assisted study of *Kulturkampf* in German, Polish and English. *Pragmatics and Society*, 9(1), 117–147. <https://doi.org/10.1075/ps.16028.jaw>
- Jaworski, A. (1993). *The power of silence: Social and pragmatic perspectives*. Sage.
- Jay, T. (1992). *Cursing in America: A psycholinguistic study of dirty language in the courts, in the movies, in the schoolyards, and on the streets*. J. Benjamins Pub. Co.
- Jay, T. (2000). *Why we curse: A neuro-psycho-social theory of speech*. Benjamins.
- Jay, T., & Janschewitz, K. (2008). The pragmatics of swearing. *Journal of Politeness Research. Language, Behaviour, Culture*, 4(2). <https://doi.org/10.1515/JPLR.2008.013>
- Jewkes, Y. (2015). *Media & crime* (3rd edition). SAGE.
- Jiang, F. (Kevin), & Hyland, K. (2015). ‘The fact that’: Stance nouns in disciplinary writing. *Discourse Studies*, 17(5), 529–550. <https://doi.org/10.1177/1461445615590719>
- Johansson, S. (2006). ‘Sometimes you wanna hate celebrities’, Tabloid readers and celebrity coverage. In S. Holmes & S. Redmond (Eds.), *Framing celebrity: New directions in celebrity culture*. Routledge.
- Johnson, W. H. (2023). Pleistocene Epoch. In *Britannica*.
<https://www.britannica.com/science/Pleistocene-Epoch>
- Jordan, M. D. (2000). *The silence of Sodom: Homosexuality in modern Catholicism*. Univ. of Chicago Press.
- Judet de La Combe, P. (Writer). (2023a). Œdipe aux pieds gonflés [Broadcast]. In *Quand les Dieux rôdaient sur la Terre*. France Inter.

<https://www.radiofrance.fr/franceinter/podcasts/quand-les-dieux-rodaiant-sur-la-terre/quand-les-dieux-rodaiant-sur-la-terre-du-samedi-25-fevrier-2023-4014199>

Judet de La Combe, P. (Writer). (2023b). Œdipe, le damné [Broadcast]. In *Quand les Dieux rôdaient sur la Terre*. France Inter.

<https://www.radiofrance.fr/franceinter/podcasts/quand-les-dieux-rodaiant-sur-la-terre/quand-les-dieux-rodaiant-sur-la-terre-du-samedi-04-mars-2023-7103750>

Justice, B., & Justice, R. (1980). *The broken taboo, Sex in the family*.

Kacprzak, A. (2018). Les regards sur la vieillesse : Le tabou et son contournement. *Acta Universitatis Lodziensis. Folia Litteraria Romanica*, 12, 57–69.

<https://doi.org/10.18778/1505-9065.12.06>

Kantor, J., & Twohey, M. (2017). Harvey Weinstein Paid Off Sexual Harassment Accusers for Decades. *The New York Times*. <https://www.nytimes.com/2017/10/05/us/harvey-weinstein-harassment-allegations.html>

Keeble, R. (2009). *Ethics for journalists* (2nd ed). Routledge.

Khairullina, R. K., Fatkullina, F. G., So, Q., & Lin, Z. (2020). *Taboo As A Linguistic And Cultural Phenomenon*. 1969–1975. <https://doi.org/10.15405/epsbs.2020.10.05.259>

Kitzinger, J. (1998). The gender-politics of news production, Silenced voices and false memories, Jenny Kitzinger. In C. Carter, G. Branston, & S. Allan (Eds.), *News, gender, and power*. Routledge.

Klemperer, V. (2006). *LTI, la langue du IIIe Reich : Carnets d'un philologue* (É. Guillot, Trans.). Michel.

- Koester, A. (2007). 'About Twelve Thousand or So': Vagueness in North American and UK Offices. In J. Cutting, *Vague Language Explored*. Palgrave Macmillan.
- Kouchner, C. (2021). *La familia grande*. Éditions du Seuil.
- Kress, G. R., & Van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. Arnold ; Oxford University Press.
- Krieg-Planque, A. (2004). Souligner l'euphémisme : Opération savante ou acte d'engagement ? Analyse du « jugement d'euphémisation » dans le discours politique. *Semen*, 17. <https://doi.org/10.4000/semn.2351>
- Kuhn, R. (2011). *The media in contemporary France*. Open University Press.
- Labasse, B. (2012). *Sexe, sang et physique des particules : Le « sensationnalisme » est-il partout... ou nulle part ?*
- L'Affaire d'Outreau*. (2023). [Broadcast]. France 2. <https://www.france.tv/france-2/l-affaire-d-outreau/>
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.
- Lamy, R. (2023). *En bons pères de famille*. JC Lattès.
- Langton, R. (2009). *Sexual Solipsism*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199247066.001.0001>
- Leach, E. R. (1982). *Social anthropology*. Fontana.
- Lévi-Strauss, C. (1949). *Nature, culture et société: Les structures élémentaires de la parenté, chapitres I et II*. Presses Universitaires de France.

- Liptak, A. (2022). In 6-to-3 Ruling, Supreme Court Ends Nearly 50 Years of Abortion Rights. *The New York Times*. <https://www.nytimes.com/2022/06/24/us/roe-wade-overturned-supreme-court.html>
- López Díaz, M. (2013). Quand dire, c'est édulcorer et occulter: L'euphémisme dans l'information médiatique. *Journal of French Language Studies*, 23(3), 377–399. <https://doi.org/10.1017/S0959269512000269>
- López Díaz, M. (2014). *Deux variantes stylistiques dans le même contexte: Le tabou et l'euphémisme. Figures du discours et contextualisation, Actes du colloque de Nice*. <http://revel.unice.fr/symposia/figuresetcontextualisation/?id=1313>
- López Díaz, M. (2018a). La cooccurrence du tabou et de l'euphémisme ou les conditions de la synonymie: *Travaux de linguistique*, n° 76(1), 27–42. <https://doi.org/10.3917/tl.076.0027>
- López Díaz, M. (2018b). La pratique de la redénomination à l'instar du politiquement correct. *Acta Universitatis Lodzensis. Folia Litteraria Romanica*, 12, 19–30. <https://doi.org/10.18778/1505-9065.12.03>
- Machin, D., & Mayr, A. (2023). *How to do critical discourse analysis: A multimodal introduction* (Second edition). SAGE Publications.
- MacMullen, R. J. (1990). *Male rape: Breaking the silence on the last taboo*. Gay Men's Pr.
- Marchi, A., & Taylor, C. (2016). If on a Winter's Night Two Researchers... A Challenge to Assumptions of Soundness of Interpretation. *Critical Approaches to Discourse Analysis Across Disciplines*, 67, 372–373. <https://doi.org/10.5209/CLAC.53494>

- Marcotte, A. (2016). *Debunking genetic sexual attraction: Incest by any other name is still incest*. <https://www.salon.com/2016/08/16/debunking-genetic-sexual-attraction-incest-by-any-other-name-is-still-incest/>
- McCarthy, M., & O’Keeffe, A. (2010). Historical Perspective: What are corpora and how have they evolved? In *The Routledge Handbook of Corpus Linguistics*.
- McEnery, A., & Brookes, G. (2022). Building a written corpus: What are the basics? In A. O’Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (Second edition). Routledge.
- McEnery, A., & Xiao, Z. (2004). Swearing in Modern British English: The Case of Fuck in the BNC. *Language and Literature: International Journal of Stylistics*, 13(3), 235–268. <https://doi.org/10.1177/0963947004044873>
- McEnery, T. (2006). *Swearing in English: Bad language, purity and power from 1586 to the present*. Routledge.
- McEnery, T., & Baker, H. (2017). *Corpus linguistics and 17th-century prostitution: Computational linguistics and history*. Bloomsbury Academic.
- McEnery, T., & Brezina, V. (2022). *Fundamental principles of corpus linguistics*. Cambridge University Press.
- McNair, B. (2009). *News and journalism in the UK* (5th ed). Routledge.
- Media Reform Coalition. (2021). *Who owns the UK media?* https://www.mediareform.org.uk/wp-content/uploads/2021/03/Who-Owns-the-UK-Media_final2.pdf

- Mehrabian, A. (1981). *Silent messages: Implicit communication of emotions and attitudes* (2. ed). Wadsworth.
- Meillet, A. (1906). *Quelques hypothèses sur des interdictions de vocabulaire dans les langues indo-européennes*.
- Montgomery, E., Pope, C., & Rogers, J. (2015). A feminist narrative study of the maternity care experiences of women who were sexually abused in childhood. *Midwifery*, 31(1), 54–60. <https://doi.org/10.1016/j.midw.2014.05.010>
- Morley, J., & Partington, A. (2009). A few *Frequently Asked Questions* about semantic—Or evaluative—Prosody. *International Journal of Corpus Linguistics*, 14(2), 139–158. <https://doi.org/10.1075/ijcl.14.2.01mor>
- Morris, D. (1986). *The naked ape* (Reprinted). Triad Grafton Books.
- Moschonas, S., & Spitzmüller, J. (2010). *Prescriptivism in and about the media: A comparative analysis of corrective practices in Greece and Germany*.
- Nardone, C. (2016). ‘Women and Work’: a Cross-Linguistic Corpus-Assisted Discourse Study in German and in Italian. *Critical Approaches to Discourse Analysis Across Disciplines*, 67, 372–373. <https://doi.org/10.5209/CLAC.53494>
- O’Driscoll, J. (2020). *Offensive language: Taboo, offence and social control*. Bloomsbury Academic.
- Office for National Statistics. (2021). *Sexual offences victim characteristics, England and Wales: Year ending March 2020*. Office for National Statistics.
- Office for National Statistics. (2024). *Families and households in the UK: 2023*.

- Office of Communications. (2023). *News consumption in the UK: 2023*.
- Office of Communications, O. (2019). *News consumption in the UK: 2019*.
https://www.ofcom.org.uk/__data/assets/pdf_file/0027/157914/uk-news-consumption-2019-report.pdf
- O’Keeffe, A., & McCarthy, M. (Eds.). (2022). *The Routledge Handbook of Corpus Linguistics* (Second edition). Routledge.
- Oliver, M. (2003). Richard Gardner, 72; Had Theory on False Claims of Abuse Against Parents. *Los Angeles Times*. <https://www.latimes.com/archives/la-xpm-2003-jun-12-me-gardner12-story.html>
- O’Neill, D. (2012). *No cause for celebration: The rise of celebrity news values in the British quality press*. 1(2).
- Ouest-France. (2024). *Inceste: Le juge Édouard Durand, ex-coprésident de la Ciivise, revient sur son éviction*. <https://www.ouest-france.fr/societe/faits-divers/inceste/inceste-le-juge-edouard-durand-ex-copresident-de-la-ciivise-revient-sur-son-eviction-2719c54a-c685-11ee-9c4f-9be0695f50c2#:~:text=En%20décembre%202023%2C%20le%20juge,de%20la%20Ciivise%20ont%20démissionné>.
- Ovidie. (2022). Petite sémiotique de la step mom. Pornographies et culture de l’inceste. In *La Culture de l’inceste* (Seuil).
- Partington, A. (2004). ‘Utterly content in each other’s company’: Semantic prosody and semantic preference. *International Journal of Corpus Linguistics*, 9(1), 131–156.
<https://doi.org/10.1075/ijcl.9.1.07par>

- Partington, A. (2014). Mind the gaps: The role of corpus linguistics in researching absences. *International Journal of Corpus Linguistics*, 19(1), 118–146.
<https://doi.org/10.1075/ijcl.19.1.05par>
- Partington, A. (2018). Intimations of ‘Spring’? What Got Said and What Didn’t Get Said about the Start of the Middle Eastern/North African Uprisings: A Corpus-assisted Discourse Study of a Historical Event. In M. Schröter & C. Taylor (Eds.), *Exploring Silence and Absence in Discourse, Empirical Approaches*. Palgrave.
- Partington, A., Duguid, A., & Taylor, C. (2013). *Patterns and meanings in discourse: Theory and practice in corpus-assisted discourse studies (CADS)*. John Benjamins Publishing Company.
- Partington, A., & Marchi, A. (2015). *Using corpora in discourse analysis*.
- Pasolini, P. P. (Director). (1967). *Edipo Re* [Video recording].
- Payne, L. (2023). Euphemism. In *Britannica*. <https://www.britannica.com/art/euphemism>
- Pedraza, A. P. (Ed.). (2018). *Linguistic taboo revisited: Novel insights from cognitive perspectives*. International Cognitive Linguistics, Berlin ; Boston. De Gruyter Mouton.
- Penfold-Mounce, R. (2009). *Celebrity culture and crime: The joy of transgression*. Palgrave Macmillan.
- Pereda, N., Guilera, G., Forns, M., & Gómez-Benito, J. (2009). The prevalence of child sexual abuse in community and student samples: A meta-analysis. *Clinical Psychology Review*, 29(4), 328–338. <https://doi.org/10.1016/j.cpr.2009.02.007>

Perilli, B. (2023). Florida, insegnante licenziata per aver mostrato il David di Michelangelo ai suoi studenti. *La Repubblica*.

https://www.repubblica.it/esteri/2023/03/24/news/professoressa_florida_david_michelangelo_pornografia-393494524/

Piazza, R. (2009). News is reporting what was said: Techniques and patterns of attribution. In L. Haarman & L. Lombardo (Eds.), *Evaluation and stance in war news: A linguistic analysis of American, British and Italian television news reporting of the 2003 Iraqi war* (pp. 170–194). Continuum.

Picq, P., & Brenot, P. (2012). *Le sexe, l'homme et l'évolution*. O. Jacob.

Piterbraut-Merx, T. (2022). Oreilles cousues et mémoires mutines. L'inceste et les rapports de pouvoir adulte-enfant. In I. Brey & J. Drouar, *La Culture de l'inceste*. Editions du Seuil.

Popescu, M. (2018). Le tabou linguistique. Un paradoxe toujours actuel. *Acta Universitatis Lodzianis. Folia Litteraria Romanica*, 12, 135–145. <https://doi.org/10.18778/1505-9065.12.13>

Porter, R., & Ciba Foundation (Eds.). (1984). *Child sexual abuse within the family*. Tavistock Publ.

Prażuch, W. (2018). Le tabou et sa dénonciation – la parole publique entre la « langue de bois » et le « parler vrai ». *Acta Universitatis Lodzianis. Folia Litteraria Romanica*, 12, 147–159. <https://doi.org/10.18778/1505-9065.12.14>

Putland, E., Chikodzore-Paterson, C., & Brookes, G. (2023). Artificial intelligence and visual discourse: A multimodal critical discourse analysis of AI-generated images of “Dementia”. *Social Semiotics*.

- Quarta, E. (2009). *Zones sensibles ou ghettos? L'euphémisme dans la représentation des banlieues françaises dans les quotidiens français et italiens.*
- Rashkow, I. N. (2000). *Taboo or not taboo: Sexuality and family in the Hebrew Bible.* Fortress Press.
- Regional Office For Africa, W. H. O. (2004). *Child Sexual Abuse: A Silent Health Emergency.*
- Richardson, J. E. (2007). *Analysing newspapers: An approach from critical discourse analysis.* Palgrave Macmillan.
- Rodríguez Pedreira, N. (2018). De la « mère porteuse » à la « GPA » : Vers un contournement du tabou. *Acta Universitatis Lodziensis. Folia Litteraria Romanica*, 12, 31–46.
<https://doi.org/10.18778/1505-9065.12.04>
- Rojek, C. (2001). *Celebrity.* Reaktion Books.
- Rojek, C. (2012). *Fame attack: The inflation of celebrity and its consequences.* Bloomsbury Academic.
- Romero, M. (2018). L'inceste et le droit pénal français contemporain : Lorsque l'application du droit dans les tribunaux révèle une indicible frontière. *GLAD!*, 05.
<https://doi.org/10.4000/glad.1230>
- Romero, M. (2020). La parole de l'enfant victime de violences sexuelles : Une enquête au sein de tribunaux correctionnels français en 2010. *Criminocorpus, Revue Hypermédia.*
<https://doi.org/10.4000/criminocorpus.7177>
- Rosewarne, L. (2013). *American taboo: The forbidden words, unspoken rules, and secret morality of popular culture.* Praeger, An Imprint of ABC-CLIO, LLC.

Rosulek, L. F. (2008). Manipulative silence and social representation in the closing arguments of a child sexual abuse case. *Text & Talk - An Interdisciplinary Journal of Language, Discourse Communication Studies*, 28(4), 529–550.

<https://doi.org/10.1515/TEXT.2008.026>

Rougier, M. (2020). *La parole de l'enfant victime de crimes sexuels dans les procédures judiciaires aux Assises de la Charente (1889-1914). L'enfance au tribunal. Enjeux historiques, perspectives contemporaines*(Criminocorpus).

Runyan, D. K. (1998). Prevalence, Risk, Sensitivity, and Specificity: A Commentary on the Epidemiology of Child Sexual Abuse and the Development of a Research Agenda. *Child Abuse & Neglect*, 22(6), 493–498. [https://doi.org/10.1016/S0145-2134\(98\)00015-5](https://doi.org/10.1016/S0145-2134(98)00015-5)

Russell, D. E. H. (1983). The incidence and prevalence of intrafamilial and extrafamilial sexual abuse of female children. *Child Abuse & Neglect*, 7(2), 133–146. [https://doi.org/10.1016/0145-2134\(83\)90065-0](https://doi.org/10.1016/0145-2134(83)90065-0)

Russell, D. E. H. (1999). *The secret trauma: Incest in the lives of girls and women ; with a new introduction by the author* (Revised edition). Basic Books.

Ryan, C., & Jethá, C. (2021). *Au commencement était le sexe: Aux origines préhistoriques de la sexualité humaine* (R. Robert, Trans.). Alisio.

Ryan, C., & Jethá, C. (with Picq, P.). (2023). *Au commencement était le sexe: Aux origines préhistoriques de la sexualité humaine* (R. Robert, Trans.). J'ai lu.

Sablayrolles, J.-F. (2017). Quelques facteurs de variation des interdits langagiers. *ACTA UNIVERSITATIS LODZIENSIS. FOLIA LITTERARIA ROMANICA*, 12, 9–18. <https://doi.org/10.18778/1505-9065.12.02>

Sabri, O., Manceau, D., & Pras, B. (2010). Le tabou, un concept peu exploré en marketing.

Recherche et Applications en Marketing, 25(1), 59–86.

<https://doi.org/10.1177/076737011002500103>

Sandoz, C. (2022). Le tabou linguistique comme facteur de processus dérivationnels. *Cahiers*

du Centre de Linguistique et des Sciences du Langage, 19, 139–146.

<https://doi.org/10.26034/la.cdclsl.2005.1562>

Sariola, H., & Uutela, A. (1996). The prevalence and context of incest abuse in Finland. *Child*

Abuse & Neglect, 20(9), 843–850. [https://doi.org/10.1016/0145-2134\(96\)00072-5](https://doi.org/10.1016/0145-2134(96)00072-5)

Saussure, F. de. (1916). *Cours de linguistique générale* (C. Bally, A. Sechehaye, & A. Riedlinger,

Eds.). Payot.

Saville-Troike, M. (1985). The Place of Silence in an Integrated Theory of Communication. In

Perspectives on Silence. Ablex Pub. Corp.

Schmid, H.-J. (2012). *English Abstract Nouns as Conceptual Shells: From Corpus to Cognition*.

De Gruyter Mouton.

Schröter, M., & Taylor, C. (2018). *Exploring Silence and Absence in Discourse, Empirical*

Approaches. Palgrave.

Schwark, S. (2017). Visual Representations of Sexual Violence in Online News Outlets.

Frontiers in Psychology, 8, 774. <https://doi.org/10.3389/fpsyg.2017.00774>

Scott, M., & Tribble, C. (2006). *Textual patterns: Key words and corpus analysis in language*

education. J. Benjamins.

- Semino, E., Demjén, Z., Hardi, A., Payne, S., & Rayson, P. (2020). *Metaphor, cancer and the end of life: A corpus-based study*. Routledge.
- Sim, S. (2007). *Manifesto for silence: Confronting the politics and culture of noise*. Edinburgh Univ. Press.
- Sinno, N. (2023). *Triste tigre*. P.O.L.
- Skidmore, P. (1998). Gender and the Agenda, News reporting of child sexual abuse. In C. Carter, G. Branston, & S. Allan (Eds.), *News, gender, and power*. Routledge.
- Sophocle. (1942). *Trachiniennes. Philoctète. Oedipe à Colone. Les Limiers*. (P. Masqueray, Ed.; 2nd ed.). Les Belles Lettres.
- Sophocle. (2015). *Oedipe roi* (J.-L. Backès & R. Dreyfus, Eds.; J. Grosjean, Trans.). Gallimard.
- Sophocles. (2020). *Oedipus Rex*. Open Road Media.
- Spencer-Oatey, H. (2012). *What is Culture? What is Culture? A Compilation of Quotations*. GlobalPAD Core Concepts.
- Staats, A. W. (1996). *Behavior and personality: Psychological behaviorism*. Springer.
- Strand, C. (2018). Cross-media Studies as a Method to Uncover Patterns of Silence and Linguistic Discrimination of Sexual Minorities in Ugandan Print Media. In M. Schröter & C. Taylor (Eds.), *Exploring Silence and Absence in Discourse* (pp 125-157). Palgrave.
- Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. Blackwell Publishers.
- Summers, H., & Campbell, B. (2022). Parental alienation and the unregulated experts shattering children's lives. *The Guardian*.

- Taboada, M. T. (Ed.). (2013). *Contrastive discourse analysis: Functional and corpus perspectives*. Equinox.
- Tannen, D., Hamilton, H. E., & Schiffrin, D. (Eds.). (2015). *The handbook of discourse analysis* (Second edition). John Wiley & Sons, Inc.
- Tannen, D., & Saville-Troike, M. (Eds.). (1985). *Perspectives on silence*. Ablex Pub. Corp.
- Taylor, M., & Ogden, J. (2005). Doctors' use of euphemisms and their impact on patients' beliefs about health: An experimental study of heart failure. *Patient Education and Counseling*, 57(3), 321–326. <https://doi.org/10.1016/j.pec.2004.09.001>
- Taylor, C. (2013). Searching for similarity using corpus-assisted discourse studies. *Corpora*, 8(1), 81–113. <https://doi.org/10.3366/cor.2013.0035>
- Taylor, C. (2014). Investigating the representation of migrants in the UK and Italian press: A cross-linguistic corpus-assisted discourse analysis. *International Journal of Corpus Linguistics*, 19(3), 368–400. <https://doi.org/10.1075/ijcl.19.3.03tay>
- Taylor, C., & Marchi, A. (Eds.). (2018). *Corpus approaches to discourse: A critical review*. Routledge.
- The New York Times. (2024). Tracking Abortion Bans Across the Country. *The New York Times*. <https://www.nytimes.com/interactive/2022/us/abortion-laws-roe-v-wade.html>
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. J. Benjamins.
- Tranchese, A. (2023). *From Fritzl to #metoo: Twelve Years of Rape Coverage in the British Press*. Springer International Publishing. <https://doi.org/10.1007/978-3-031-09353-1>

- Trinch, S. L. (2001). Managing euphemism and transcending taboos: Negotiating the meaning of sexual assault in Latinas narratives of domestic violence. *Text - Interdisciplinary Journal for the Study of Discourse*, 21(4). <https://doi.org/10.1515/text.2001.012>
- Tuggy, D. (1993). Ambiguity, polysemy, and vagueness. *Cogl*, 4(3), 273–290. <https://doi.org/10.1515/cogl.1993.4.3.273>
- Turner, G. (1999). Tabloidization, journalism and the possibility of critique. *International Journal of Cultural Studies*, 2(1), 59–76. <https://doi.org/10.1177/136787799900200104>
- Turner, G. (2014a). *Is celebrity news, news?* *Journalism* (2014, Vol. 15(2) 144–152). <https://doi.org/10.1177/1464884913488719>
- Turner, G. (2014b). *Understanding celebrity* (2. ed). SAGE.
- Turner, G. (2016). Celebrity, Participation, and the Public. In P. D. Marshall & S. Redmond (Eds.), *A companion to celebrity*. Wiley Blackwell.
- Vallery, R., & Lemmens, M. (2021). The sound of taboo: Exploring a sound-meaning association in swear words of English and French. *Pragmatics & Cognition*, 28(1), 87–137. <https://doi.org/10.1075/pc.20021.val>
- Van Den Bulck, H., Paulussen, S., & Bels, A. (2017). Celebrity news as hybrid journalism: An assessment of celebrity coverage in Flemish newspapers and magazines. *Journalism*, 18(1), 44–63. <https://doi.org/10.1177/1464884916657523>
- Van Dijk, T. A. (2001). Critical Discourse Analysis. In D. Schiffrin, D. Tannen, & H. E. Hamilton (Eds.), *The handbook of discourse analysis*. Blackwell Publishers.

- Van Dijk, T. A. (2006). Ideology and discourse analysis. *Journal of Political Ideologies*, 11(2), 115–140. <https://doi.org/10.1080/13569310600687908>
- Van Leeuwen, T. (1996). The representation of social actors. In C. R. Caldas-Coulthard & M. Coulthard, *Texts and Practices: Readings in Critical Discourse Analysis* (1st ed). Taylor and Francis.
- Van Leeuwen, T. (2008). *Discourse and practice: New tools for critical discourse analysis*. Oxford University Press.
- van Rooij, R. (2011). Vagueness and Linguistics. In G. Ronzitti (Ed.), *Vagueness: A Guide* (pp. 123–170). Springer Netherlands. https://doi.org/10.1007/978-94-007-0375-9_6
- Verdrager, P. (2021). *Le grand renversement: Pédocriminalité: comment en est-on arrivé là?* Armand Colin.
- Vessey, R. (2013). Challenges in cross-linguistic corpus-assisted discourse studies. *Corpora*, 8(1), 1–26. <https://doi.org/10.3366/cor.2013.0032>
- Vessey, R. (2024). *The language of sexual violence and impropriety*.
- Vinel, V. (2017). En famille: Une libération de la parole sur le corps sexuel ? In N. Diasio & V. Vinel (Eds.), *Corps et préadolescence* (pp. 61–78). Presses universitaires de Rennes. <https://doi.org/10.4000/books.pur.148800>
- Vinterberg, T. (Director). (1998). *The Celebration* [Video recording].
- Webster, H. (1973). *Taboo: A sociological study* (Reprint of 1942). Octagon Books.
- Widłak, S. (1965). L'interdiction linguistique en français d'aujourd'hui. *Revue belge de philologie et d'histoire*, 43(3), 932–945. <https://doi.org/10.3406/rbph.1965.2587>

Wodak, R. (2014). Critical Discourse Analysis. In C. Leung & B. V. Street (Eds.), *The Routledge handbook of English studies*. Routledge, Taylor & Francis Group.

Wodak, R., & Meyer, M. (Eds.). (2016). *Methods of critical discourse studies* (3rd edition). SAGE.

Wolf, A. P., & Durham, W. H. (Eds.). (2005). *Inbreeding, incest, and the incest taboo: The state of knowledge at the turn of the century*. Stanford University Press.

Wykes, M. (1998). A family affair, The British press, sex and the Wests. In C. Carter, G. Branston, & S. Allan (Eds.), *News, gender, and power*. Routledge.

Xiao, R., & McEnery, T. (2014). *Corpus-based contrastive studies of English and Chinese*. Routledge.

Yildirim, A., Ozer, E., Bozkurt, H., Ozsoy, S., Enginyurt, O., Evcuman, D., Yilmaz, R., & Kuyucu, Y. E. (2014). Evaluation of social and demographic characteristics of incest cases in a university hospital in Turkey. *Medical Science Monitor*, 20, 693–697.
<https://doi.org/10.12659/MSM.890361>

Zerubavel, E. (2008). *The elephant in the room: Silence and denial in everyday life*. Oxford University Press.