



HAL
open science

Pour une étude de la terminologie médicale de Proust : rétro-numérisation et analyse de la Correspondance avec sa mère

Federica Vezzani, Ludovico Monaci

► **To cite this version:**

Federica Vezzani, Ludovico Monaci. Pour une étude de la terminologie médicale de Proust : rétro-numérisation et analyse de la Correspondance avec sa mère. *L'analisi linguistica e letteraria*, 2022, XXX (2), pp.127-140. <hal-04618807>

HAL Id: hal-04618807

<https://hal.science/hal-04618807v1>

Submitted on 25 Jun 2024

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

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



HAL Authorization

Pour une étude de la terminologie médicale de Proust : rétro-numérisation et analyse de la *Correspondance avec sa mère**

Ludovico Monaci

ludovico.monaci@phd.unipd.it

Università degli Studi di Padova.

Federica Vezzani

federico.vezzani@unipd.it

Università degli Studi di Padova.

This paper aims to highlight the potential of a terminological analysis approach to literary texts. We present a work of retro-digitization and terminological study of a corpus of letters exchanged between Proust and his mother. The identified medical terminology constitutes the object of investigation to illustrate how specialized lexical units emigrate from intimate writings to join the experimental laboratory of the novel. The illustrated case study will focus on the term “trional” and its evolution from the *Correspondance* to the *Recherche*.

Keywords: *Digital Humanities, Medical Terminology, Marcel Proust, Correspondance, Textual Genesis.*

*La médecine ne prouve-t-elle pas que la
chronologie n'a rien à voir avec l'âge ?¹*

Introduction

L'échange épistolaire entre Marcel Proust (1871-1922) et sa mère Jeanne Weil (1849-1905) restitue l'image d'une femme prête à écouter et à soulager (mais aussi à accentuer et à alimenter) les malaises de son fils. Comme de nombreuses études ont été consacrées à ce sujet, notre but n'est pas d'enquêter sur le rapport à la fois symbiotique et ambivalent qui liait l'écrivain à sa mère, bien que, certes, le rôle d'interlocutrice privilégiée que Jeanne Weil joue auprès de son fils malade représente le point de départ heuristique de cette enquête. À travers une approche terminologique axée sur un corpus numérisé au préalable, cette contribution vise à mettre au jour comment des termes simples (un seul mot-forme) et des termes complexes (plusieurs mots-formes) relevant du domaine médical émigrent des écrits intimes échangés entre l'auteur et sa mère pour rejoindre le laboratoire expérimental du roman. En outre, nous proposons de fournir un modèle opérationnel

* Les deux auteurs ont contribué, dans la même mesure, au cadre méthodologique et à l'approche utilisés, d'une part, et, d'autre part, à l'interprétation qualitative des données. Ludovico Monaci a rédigé les sections 1 et 3, Federica Vezzani la section 2. L'introduction et les conclusions ont été conçues et rédigées conjointement.

¹ M. Proust, *Correspondance*, P. Kolb ed., Plon, Paris 1970-1993, vol. XVI, p. 204 (dorénavant abrégé *Corr.*, suivi du tome en chiffres romains et du numéro de page en chiffres arabes).

pouvant être étendu – dans une perspective synchronique ou diachronique – soit à l'ensemble du corpus des lettres soit à d'autres domaines spécifiques.

La structure de l'article est la suivante. Dans la première partie, nous résumerons les études qui ont été menées jusqu'ici au sujet de l'influence que la médecine et les médecins ont exercée sur Proust. La deuxième partie du travail est consacrée aux phases de préparation, de traitement et de gestion du corpus : nous décrirons la méthodologie adoptée afin de rétro-numériser la correspondance et d'en extraire la terminologie médicale pertinente. Enfin, l'étude de cas autour du terme simple « trional » témoignera de la valeur ajoutée d'une approche terminologique appliquée à un corpus de lettres.

1. La *Recherche* médicale de Marcel Proust

Selon le témoignage de son frère cadet², Marcel Proust est victime de sa première crise d'asthme à l'âge de 9 ans : cet élément suffit pour affirmer que l'écrivain est un malade avant d'être un hypocondriaque. À côté de cette évidence biographique, l'auteur de la *Recherche* possède une profonde connaissance médicale. Ayant grandi dans une famille de médecins (son père Adrien est un hygiéniste épidémiologiste ; son frère Robert est un chirurgien gynécologue et urologue), Marcel donnait des conseils de santé à ses amis et pratiquait l'automédication. Il est évident que la maladie, la quête spasmodique (et vaine) d'une cure pour y remédier, ainsi que la culture médicale sont des vecteurs importants autour desquels la biographie, la correspondance et l'œuvre de l'auteur doivent être lues.

Au début de mai 1908, au beau milieu de la composition des « soixante-quinze feuillets »³, Marcel Proust renseigne Louis d'Albufera sur les projets qu'il a en chantier :

Car j'ai en train : une étude sur la noblesse / un roman parisien / un essai sur Sainte-Beuve et Flaubert / un essai sur les Femmes / un essai sur la Pédérastie / (pas facile à publier) / une étude sur les vitraux / une étude sur les pierres tombales / une étude sur le roman⁴.

L'essai sur Flaubert sera publié dans la *NRF*⁵, alors que les autres projets convergeront plus ou moins variablement vers le dénominateur commun qu'est la *Recherche*. Dans la pléiade des œuvres en puissance ne figure pas une étude sur le discours médical. Pourtant, en août 1905, un mois avant la mort de sa mère, Proust écrit à Anna de Noailles : « Je vais faire un livre sur les médecins »⁶. De plus, en 1908, quelques jours avant de dresser la liste des projets « en train », il révèle à Maurice de Fleury : « Je m'étais amusé quelquefois à faire des pastiches de nature médicale ! »⁷. Le discours scientifique a tellement animé l'existence de l'écrivain que l'idée latente d'en faire l'objet d'un ouvrage littéraire devance même la mise en jeu des instances et des éléments qui ont jeté les bases de la cathédrale de la *Recherche*. Dans cet esprit, Cottard, du Boulbon, Percepied, Dieulafoy et le Professeur

² R. Proust, *Marcel Proust intime*, « La Nouvelle Revue Française. Hommage à Marcel Proust », 112, 1923, pp. 24-26.

³ Il s'agit du manuscrit qui est censé être le premier noyau narratif de la *Recherche* : « La rédaction [...] s'est échelonnée entre les premiers mois et l'automne de 1908 ; leur élaboration a peut-être commencé dès la fin de 1907 » (M. Proust, *Les soixante-quinze feuillets et autres manuscrit inédits*, N. Mauriac Dyer ed., Gallimard, Paris 2021, p. 18).

⁴ *Corr.*, VIII, pp. 112-113.

⁵ M. Proust, *Essais et articles*, P. Clarac – Y. Sandre ed., Gallimard, Paris 1971 (Bibliothèque de la Pléiade, 229), pp. 586-600.

⁶ *Corr.*, V, p. 318.

⁷ *Corr.*, VIII, p. 74.

E*** concentrent les qualités et les défauts des spécialistes que Proust a consultés le long de sa vie⁸.

Beaucoup d'études et de colloques ont été consacrés à l'influence que le discours médical a exercée sur les univers sémiotique et stylistique de la *Recherche*⁹, ainsi qu'aux intuitions littéraires qui pourraient avoir devancé des découvertes scientifiques : les enquêtes font appel à la médecine au sens large¹⁰, à la psychologie¹¹, mais aussi aux neurosciences¹². En tenant compte de l'avancement du projet *Corr-Proust*¹³, nous proposons une analyse ciblée sur une collection de lettres de Marcel Proust par le biais d'une approche terminologique qui privilégie l'analyse des occurrences : le support de l'extraction terminologique jettera une lumière nouvelle sur la correspondance de l'auteur et sur la façon dont elle intervient dans les différentes étapes génétiques d'un épisode romanesque et dans l'ensemble de l'œuvre.

2. Rétro-numérisation, traitement et gestion du corpus

Au vu de la forte cohésion de l'ensemble des documents, nous avons choisi d'adopter comme corpus d'analyse les lettres qui, publiées en 1953 par Philip Kolb¹⁴, ont été ensuite intégrées à l'édition de la correspondance générale (1970-1993). Il s'agit d'un recueil de 149 lettres, écrites entre 1887 et 1905 et présentées par ordre chronologique : 57 ont été envoyées par Jeanne Weil à son fils ; 89 par Marcel Proust à sa mère ; 2 par Marcel Proust à son père et une par le frère de Marcel à sa mère.

Une fois le champ d'enquête circonscrit, nous décrivons les étapes de mise en forme et d'interrogation du corpus qui ont été réalisées afin d'analyser qualitativement la terminologie médicale contenue dans cette correspondance. En particulier, nous nous concentrons sur :

- 1) le flux de travail OCR¹⁵ pour la rétro-numérisation du texte,
- 2) l'annotation morphosyntaxique du corpus ainsi constitué,
- 3) l'extraction des termes médicaux pertinents pour l'étude proposée.

Les lettres collectées dans l'édition utilisée ici sont disponibles seulement au format papier. Afin d'exploiter le potentiel des systèmes de traitement et de gestion automatiques du

⁸ J. Bogousslavsky, *Marcel Proust's Lifelong Tour of the Parisian Neurological Intelligentsia: From Brissaud and Dejerine to Sollier and Babinski*, « European Neurology », 57, 2007, pp. 129-136.

⁹ *Proust's Recherche und die Medizin*, M. Föcking ed., Insel, Berlin 2014 ; *Littérature et médecine. Le cas de Proust*, M. Naturel ed., Hermann, Paris 2018 ; G. Rivane, *Influence de l'asthme sur l'œuvre de Proust*, La Nouvelle Édition, Paris 1945 ; S. Béhar, *L'univers médical de Marcel Proust*, Gallimard, Paris, 1970 ; J. Milly, *Le style de la maladie chez Proust*, « Bulletin Marcel Proust », 43, 1993, pp. 58-71. Par souci de complétude, il faut également mentionner l'étude de Brunet qui, tout en ne relevant pas du domaine médical, se configure comme la première enquête terminologique de l'œuvre de Proust : É. Brunet, *Le vocabulaire de Proust*, Honoré Champion, Paris 1983.

¹⁰ D. Wright, *Du discours médical dans À la recherche du temps perdu. Science et souffrance*, Honoré Champion, Paris 2007 ; J.-P. Ollivier, *Proust cardiologue*, Honoré Champion, Paris 2016.

¹¹ J. Yoshida, *La maladie nerveuse chez Proust : genèse du portrait du docteur du Boulbon*, « Bulletin Marcel Proust », 42, 1992, pp. 43-62 ; E. Bizub, *Proust et le moi divisé. La Recherche : creuset de la psychologie expérimentale (1874-1914)*, Droz, Genève 2006.

¹² J. Lehrer, *Proust Was a Neuroscientist*, Houghton Mifflin Harcourt, Boston 2007 ; I. Serça ed., *Proust et le temps. Un dictionnaire*, Pommier, Paris 2022.

¹³ Piloté par F. Leriche (Université Grenoble Alpes), *Corr-Proust* est le projet d'édition numérique de la correspondance de Proust <http://proust.elan-numerique.fr/> (dernière consultation le 26 mars 2022).

¹⁴ M. Proust, *Correspondance avec sa mère*, P. Kolb ed., Plon, Paris 1953.

¹⁵ *Optical Character Recognition* (en français Reconnaissance Optique des Caractères : ROC).

corpus, la première étape de notre travail a porté sur la rétro-numérisation de la collection textuelle¹⁶.

2.1. Flux de travail OCR

La méthodologie que nous avons adoptée pour la mise en forme du corpus est celle décrite dans l'étude de Del Fante et Di Nunzio¹⁷ basée, quant à elle, sur la proposition de Reul *et alii*¹⁸. Le flux de travail OCR a été divisé en trois phases fondamentales, à savoir le 1) pré-traitement, 2) le traitement et 3) le post-traitement du texte.

Dans la première phase de scannage de l'édition, nous avons procédé à la sélection des parties du texte à l'étude. Puisque la correspondance est notre seul objet d'investigation, nous avons isolé le contenu des lettres, en partant des formules d'ouverture jusqu'aux formules de clôture, et supprimé le matériel textuel non pertinent, comme les notes de commentaire de l'éditeur. Cette phase nous a permis d'obtenir des fichiers PDF qui ont ensuite été traités via le logiciel OCR *Wondershare PDFelement* 6¹⁹. L'outil nous a permis de convertir efficacement des documents scannés en texte numérique modifiable. Cependant, une phase de post-traitement du texte a été nécessaire afin de le nettoyer d'éventuelles erreurs de conversion. Le tableau 1 ci-dessous illustre les erreurs les plus fréquentes sur lesquelles nous sommes intervenus manuellement, et qui portent principalement sur une identification incorrecte d'accents, une substitution erronée de caractères, et une absence d'identification des signes de ponctuation.

Type d'erreur	Correction effectuée
Maitre	Maître
Raphael	Raphaël
Hâtait	Hâtait
Onde	Oncle

¹⁶ Une remarque de nature méta-terminologique s'impose à cet égard. 'Numériser' signifie représenter des informations données (textes, dans notre cas) sous forme de nombres (F. Vezzani, *Terminologie numérique : conception, représentation et gestion*, Peter Lang, Bern 2022). Selon l'objet d'étude et sa nature, la notion de 'numérisation' prend un double sens correspondant à un double parcours procédural. Le premier sens du terme renvoie au processus de migration des documents textuels disponibles au format papier vers des documents encodés au format numérique. Le deuxième sens du terme prend place plutôt dans les contextes qui prévoient le traitement direct d'informations sous forme numérique. La notion de 'rétro-numérisation' s'inscrit donc dans le premier sens du terme et implique l'ensemble des opérations nécessaires à la transformation en une version numérisée d'un document donné, nativement disponible dans sa version papier. À ce propos, la reconnaissance optique des caractères se configure comme une technologie pivot dans le monde des humanités numériques : elle permet d'effectuer de manière (semi-)automatisée tous les processus nécessaires à la numérisation (ou rétro-numérisation) des textes (T. Nguyen – T.T. Hai – A. Jatowt – N.-V. Nguyen – M. Coustaty – A. Doucet, *Neural Machine Translation with BERT for Post-OCR Error Detection and Correction*, « Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2020 », USA: Association for Computing Machinery, New York 2020, pp. 333-336. <https://doi.org/10.1145/3383583.3398605>; dernière consultation le 29 mars 2022).

¹⁷ D. Del Fante – G.M. Di Nunzio, *OCR Correction for Corpus-assisted Discourse Studies: A Case Study of Old Newspapers*, « Umanistica Digitale », 11, 2021, pp. 99-124 <https://doi.org/10.6092/issn.2532-8816/13689> (dernière consultation le 29 mars 2022).

¹⁸ C. Reul – D. Christ – A. Hartelt – N. Balbach – M. Wehner – U. Springmann – C. Wick – C. Grundig – A. Büttner – F. Puppe, *OCR4all—An Open-Source Tool Providing a (Semi-)Automatic OCR Workflow for Historical Printings*, « Applied Sciences », 9, 2019, 4853. <https://doi.org/10.3390/app9224853> (dernière consultation le 29 mars 2022).

¹⁹ <https://pdf.wondershare.com>

Berbe	Herbe
larmes	larmes !
Rèvant	Rêvant

Tableau 1 - Correction OCR

Une fois la phase de post-traitement de la collection textuelle terminée, les lettres ont été enregistrées sous forme de documents individuels au format Word et organisées selon la logique suivante : à chaque missive on a fait correspondre un fichier de texte, et à chaque fichier on a attribué le sigle « lettre », séparé à son tour de l'initiale du prénom du destinataire (« M » pour Marcel ; « J » pour sa mère). Par exemple, pour la première lettre de notre corpus, nous avons : « lettre_1_M ». Cette classification des textes a été adoptée afin de faciliter, dans les étapes suivantes, l'identification des contextes d'occurrence de la terminologie médicale extraite.

2.2. Annotation et gestion du corpus

Après avoir rassemblé et traité la collection, nous avons procédé aux phases de tokenisation, c'est-à-dire le découpage du texte en jetons (unités minimales qui composent le corpus de type mot), et d'annotation du corpus. Cette dernière comporte essentiellement les étapes de :

- 1) lemmatisation, à savoir le processus d'attribution du lemme à chaque mot contenu dans le corpus ;
- 2) étiquetage morphosyntaxique, c'est-à-dire le processus d'attribution à chaque jeton d'une étiquette portant des informations sur la partie du discours (nom, verbe, adjectif, etc.) correspondante.

Les deux processus décrits ci-dessus ont été réalisés automatiquement à l'aide du logiciel de gestion de corpus *Sketch Engine*²⁰. Le système disponible en ligne est doté d'un nombre considérable de fonctionnalités qui le distinguent des autres outils de gestion de corpus disponibles sur le marché²¹. Le corpus nommé « Proust – Correspondance avec sa mère : 1887-1905 » a donc été importé dans le logiciel et présente les données quantitatives illustrées dans le tableau 2 :

Jetons	Mots	Phrases	Documents
70,519	61,563	2,906	149

Tableau 2 - Informations quantitatives sur le corpus

2.3. Extraction terminologique

Parmi les différentes fonctionnalités de *Sketch Engine*, l'extraction terminologique – nommée sur le logiciel *Keywords and term extraction*²² – joue un rôle clé dans notre étude. D'un point de vue méthodologique, cette fonction repose sur une analyse contrastive de deux corpus distincts :

²⁰ <https://www.sketchengine.eu>

²¹ A. Kilgarriff – V. Baisa – J. Bušta – M. Jakubíček – V. Kovár – J. Michelfeit – P. Rychl'y – V. Suchomel, *The Sketch Engine : ten years on*, « Lexicography », 1, 2014, pp. 7-36.

²² <https://www.sketchengine.eu/guide/keywords-and-term-extraction/> (dernière consultation le 29 mars 2022).

1) le *focus corpus*, c'est-à-dire le corpus étudié (dans notre cas « Proust – Correspondance avec sa mère : 1887-1905 ») dont les termes sont extraits ; et

2) le *reference corpus*, utilisé comme corpus de contraste, généralement d'une plus grande taille que le *focus corpus* et fourni par défaut par le système (dans notre cas « French Web 2017 (frTenTen17) »).

L'opération effectuée automatiquement par le logiciel est basée sur la formule *simple maths* utilisée pour calculer le degré de « termicité » d'une séquence de caractères extraite²³. La formule identifie donc les séquences qui apparaissent plus fréquemment dans le *focus corpus* que dans le *reference corpus*. Les fréquences relatives sont employées permettant ainsi de comparer des corpus de tailles inégales. Le résultat de l'extraction automatique fournit une liste de 'candidats termes' organisés en deux pages d'affichage portant respectivement sur les candidats termes simples (constitués d'un seul mot-forme) et les candidats termes complexes (constitués de plusieurs mots-formes). La figure 1 représente une capture d'écran de la liste obtenue automatiquement pour les unités lexicales constituées d'un seul mot.

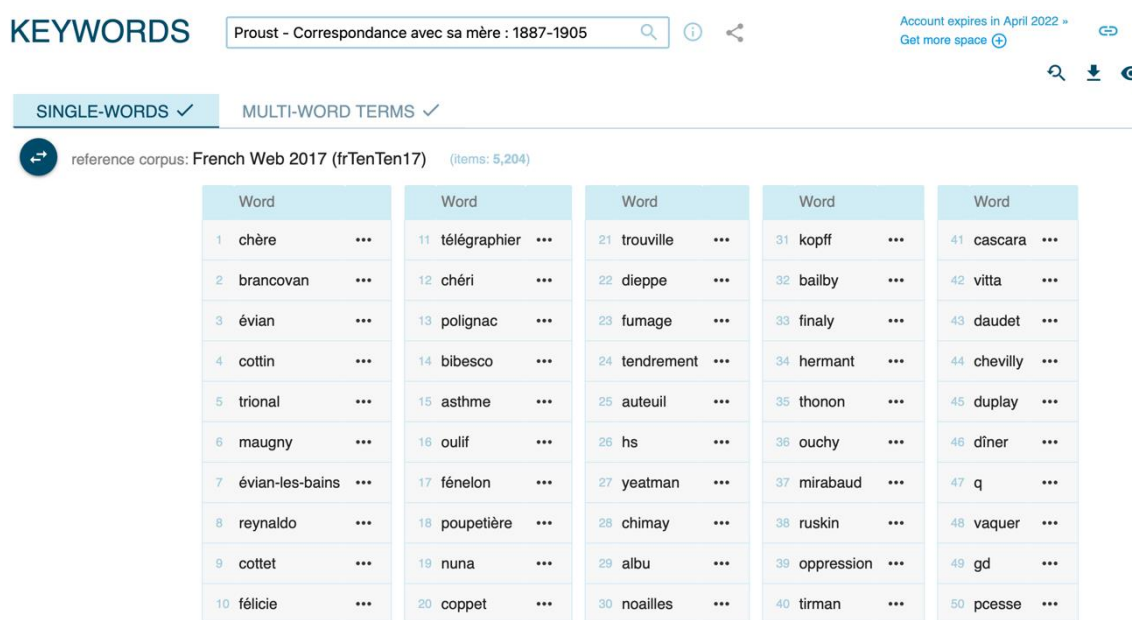


Figure 1 - *Keywords* Proust - Correspondance avec sa mère : 1887-1905

Comme on peut le constater, la liste des séquences de caractères extraites doit passer par une phase de nettoyage et de validation. Les candidats termes ont donc été analysés manuellement dans le but de sélectionner uniquement les termes pertinents au domaine médical. Pour des raisons de place, nous fournissons dans le tableau 3 ci-dessous une liste partielle des occurrences retenues aux fins de cette étude avec une indication de la fréquence absolue d'apparition au sein du *focus corpus*.

Type	Occurrence	Fréquence
Terme simple	« trional »	18
Terme simple	« asthme »	53

²³ K. Kageura – B. Umino, *Methods of automatic term recognition: A review*, « Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication » 3, 1996, pp. 259-289.

Terme simple	« fatigue »	22
Terme simple	« suffocation »	3
Terme complexe	« crise d'asthme »	8
Terme complexe	« iode morphinée »	1

Tableau 3 - Liste partielle des termes médicaux extraits

Afin de documenter la terminologie médicale utilisée, chaque occurrence fait actuellement l'objet d'une analyse qualitative. Nous avons choisi ici de nous centrer sur l'étude de cas du terme simple « trional », en montrant comment cette notion évolue et migre de la correspondance au roman.

3. Une étude de cas : le trional de la correspondance à la *Recherche*

Au sein de « cet énorme ensemble protéiforme et hétérogène »²⁴ qu'est la *Correspondance*, les lettres entre Marcel Proust et Jeanne Weil offrent un échantillon particulièrement intéressant des connaissances médicales de l'auteur. Pour répondre à l'objectif présenté *supra*, notre enquête confrontera ce corpus de la correspondance avec un épisode de l'œuvre romanesque. Tout d'abord, on conduira une analyse qualitative des occurrences du terme médical « trional » dans le corpus numérisé : le but est de mettre en valeur le jugement que les deux correspondants portent sur ce médicament. Ensuite, en s'appuyant sur les différentes déclinaisons du « baiser du soir », il s'agira de formuler des hypothèses sur la manière dont le roman intègre, absorbe (et élimine) la matière médicale, pour la distribuer ailleurs et pour en redéfinir les champs sémantiques.

3.1. Le trional dans la Correspondance avec sa mère

Comme l'illustre la figure 1, parmi les 50 premières occurrences extraites de type candidats termes simple, le terme « trional » est le premier item relevant du domaine médical. La capture d'écran présentée dans la figure 2 montre, au moyen de la fonction *Concordance*²⁵, toutes les occurrences (18 au total) du terme pivot à l'étude avec des portions de contexte.

²⁴ F. Leriche – É. Greslou, *Des éditions imprimées à l'édition numérique de la Correspondance. Enjeux éditoriaux, objectifs scientifiques, solutions d'encodage*, « Bulletin d'informations proustiennes », 49, 2019, p. 44.

²⁵ <https://www.sketchengine.eu/guide/concordance-a-tool-to-search-a-corpus/>

simple trional • 18
255.25 per million tokens • 0.026%

Left context KWIC Right context

1	□	🔍	doc#1	se pas trop de bruit. </s></s> Voilà déjà assez longtemps que je n'ai pris de	trional	et si je pouvais continuer ainsi ce serait très bien. </s></s> Je propose dans
2	□	🔍	doc#17	ontente de t'embrasser car je m'endors à la seconde même. </s></s> Sans	trional	donc pas de bruit. </s></s> Ni de fenêtres. </s></s> Je suis déjà endormi q
3	□	🔍	doc#22	</s></s> Impossibilité de me déshabiller, parce que tout froid, humide, etc.	Trional	vers 3 h et alors essay de dormir. </s></s> Je m'étends de 3 h 1/2 à 5 h, hal
4	□	🔍	doc#50	uchant et avant de me déshabiller moment critique j'ai avalé un cachet de	trional	(qui va peut-être troubler ma digestion etc). </s></s> Qu'on mette très soign
5	□	🔍	doc#88	dis qu'un simple bonjour. </s></s> Je continue à me passer rigoureuset de	trional	, amyle et valériane. </s></s> Mon dîner dans l'île avec Noufflard m'a si bie
6	□	🔍	doc#90	rouer, réveillé si tôt par les peintres que j'ai dû prendre un cachet (0,80) de	trional	, ne pouvant continuer à dormir aussi peu. </s></s> Et j'allais te demander I
7	□	🔍	doc#91	t mot. </s></s> J'ai l'ennui de te dire que j'ai pris hier soir amyle et ce matin	trional	mais cela ne durera pas. </s></s> Je vais vraiment beaucoup mieux et il m
8	□	🔍	doc#96	ment] ce qu'aura été cette nuit et si tu as pu rompre tout pacte avec l'impie	trional	. </s></s> Sandford et Merton avaient expédié ton pardessus dès hier soir.
9	□	🔍	doc#97	t. </s></s> J'ai si peur de fâcher Léon Daudet. </s></s> 3° P. S. Non pas de	Trional	. </s></s> 4° P. S. Brissaud qui connaît si bien ce pays nous eut comparé N
10	□	🔍	doc#99	papa, au moins. </s></s> Mille baisers Marcel. </s></s> Je n'ai plus pris de	trional	, car n'ayant pas de consolation toute prête je ne veux pas me payer le lux
11	□	🔍	doc#108	infimes râles ce matin. </s></s> C'est tout à fait passé. [...] Toujours pas de	trional	etc. On disait que le Matin tournait et M. de Polignac qui paraît m'apprécier
12	□	🔍	doc#109	esse Brancovan, j'ai pris ce matin pour ne pas trop refumer etc, un peu de	trional	qui a été suivi d'un sommeil réparateur et m'a très bien réussi, ce qui n'arr
13	□	🔍	doc#114	s. </s></s> Mille tendres baisers. </s></s> Marcel. </s></s> Toujours pas de	trional	(c'est sous entendu) Dépêche de Bailby : << mille pardons, dépêche égaré
14	□	🔍	doc#140	it bromure aussi, mais voudrait que je prenne tous les jours un gramme de	trional	et arrête deux jours au bout de trois jours. </s></s> Mais je préfère m'en pa
15	□	🔍	doc#142	her vers 3 heures, ou même plus tôt (de la nuit) et pour ne pas prendre de	trional	. </s></s> J'ai dormi par bribes mais enfin bien (sauf asthme qui me restait)
16	□	🔍	doc#142	it de mes nuits. </s></s> J'ai continué, encore calme, à ne plus prendre de	trional	, et j'ai refait, demi heures par demi heures [sic], une nuit suffi - sante et j'ai
17	□	🔍	doc#143	r Vaquez, qui malgré ce que je lui avais dit des inconvénients multiples du	trional	pour moi, m'avait dit de m'y mettre résolument, trouvant qu'il fallait avant to
18	□	🔍	doc#146	uner et tu m'en as instantanément un en faisant. dès que j'ai eu pris mon	trional	... clouer et crier toute la journée. </s></s> J'étais par la faute dans un tel éta

Figure 2 - Concordance « trional » Proust – Correspondance avec sa mère : 1887-1905

Le trional est un composant aliphatique²⁶ employé comme somnifère. Dans notre corpus, ce sédatif est mentionné pour la première fois en septembre 1896 (lettre_45_M) : c'est dans cette même année que le professeur Édouard Brissaud (1852-1909) publie *L'Hygiène des asthmatiques* (préfacé par Adrien Proust) et prescrit le trional à Marcel²⁷. Une seule occurrence de « trional » apparaît sous la plume de Jeanne Weil : le 20 octobre 1896, la mère demande à Marcel s'il a « pu rompre son pacte avec l'impie trional » (lettre_52_J). Dès lors, l'écrivain met à jour son « bulletin médical » : 9 fois il déclare ne pas avoir cédé à la tentation, 6 fois il admet en avoir eu recours, tandis qu'une fois seulement il reconnaît les « inconvénients multiples du trional » (lettre_95_M). Au fil du temps, ce 'sursis à statuer' évolue en plaidoyer. Dans une dédicace autographe à la marquise de Ludre, l'auteur va jusqu'à défendre le trional et le véronal²⁸ :

Je reconnais que ma manière d'user du véronal est absurde. Mais c'est un médicament utile en tant qu'entraîneur, par le sommeil artificiel préalable au sommeil naturel. Demandez à Bergson s'il ne me sait pas gré du trional²⁹.

Or, si à partir des résultats de l'extraction terminologique on repère des informations générales sur les habitudes médicales de Marcel Proust, lorsqu'on accède au roman et à son labyrinthe génétique, il est possible d'évaluer la germination et la reconversion d'éléments médicaux dans l'écriture littéraire.

²⁶ « Se dit de corps chimiques à chaîne carbonée linéaire ou ramifiée ouverte » (« Aliphatique », *Dictionnaire français de médecine et de biologie*, Paris, Masson, t. 1, 1970).

²⁷ F.-B. Michel, *Le professeur Marcel Proust*, Gallimard, Paris 2016, p. 55. Henri Vaquez (1860-1936) convient avec Brissaud de prescrire le trional à Proust (*Ibid.*, p. 101).

²⁸ À partir de 1910, Proust remplace le trional par le véronal : « Ces coups de marteau représentent la nécessité quotidienne de véronal, d'opium etc. » (*Corr.*, X, p. 51).

²⁹ *Corr.*, XVIII, p. 274. Pour une analyse pragmatique des dédicaces de Proust, on renvoie à G. Henrot Sostero, *Épistolarité apéritive. La dédicace d'exemplaire*, in *Revue d'études proustiennes*, I. Vidotto ed., 14, 2021, pp. 225-245.

3.2. Le « cachet de trional » : de la *Correspondance* aux avant-textes du « baiser du soir »

Introduit dans *Jean Santeuil*, repris dans les « soixante-quinze feuillets » et remanié dans les cahiers de brouillon avant d'aboutir à la scène des premières pages de la *Recherche*, le « baiser du soir » met en jeu une série d'éléments médicaux qui varient d'une rédaction à l'autre. À ce propos, notre *focus corpus* épistolaire s'avère une source importante de comparaison. Dans *Jean Santeuil*, où l'empreinte autobiographique est manifeste, le protagoniste attend que le docteur Surlande prenne congé de ses parents pour pouvoir recevoir le baiser de sa mère :

Ce baiser-là, c'était le viatique [...], la douce offrande de gâteaux que les Grecs attachaient au cou de l'épouse ou de l'ami défunt [...]. Ainsi Jean goûtait longuement les joues tendres de sa mère, puis sur son front fiévreux elle posait un baiser frais comme une compresse [...]³⁰.

L'imaginaire de la scène s'articule autour de deux pôles thématiques : la religion³¹ et la médecine. La première, déclinée dans l'acception païenne, s'incarne dans les libations rituelles que les Grecs anciens déposaient dans les tombes de leurs proches. En revanche, le domaine médical est représenté par le comparant³² du baiser, la compresse. Cette comparaison sera reprise dans les « soixante-quinze feuillets », qui s'appuient eux aussi sur la double référence à la religion et à la médecine pour mettre en scène le baiser du soir :

[...] et qu'enfin ce baiser précieux, unique, [...] je puisse bien en garder le souvenir entier, [...] de façon à pouvoir dans ma chambre, quand je commencerais à haleter de me sentir seul et séparé d'elle, en ouvrir le souvenir intact et gardé par mon intelligence à sa portée comme une hostie où je trouverais sa chair et son sang, ou plutôt c'était à une des modernes hosties de la science qu'il ressemblait [...], car je le rompais et le portais à mes lèvres pour qu'elles crussent retrouver la douceur de sa joue, et comme dans un cachet de narcotique j'y trouvais le sommeil³³.

L'élément alimentaire associé à l'isotopie religieuse abandonne l'imaginaire païen. La caractérisation chrétienne accentue le détournement blasphématoire de la scène : la joue condense dans une particule la chair et le sang de la mère ; c'est dans cette hostie profane et profanée que le protagoniste insomniaque trouve le sommeil « comme dans un cachet de narcotique »³⁴. Les « modernes hosties de la science » désacralisent la religion et subliment la médecine. Or, une interrogation du *focus corpus* à partir de « cachet » et de « narcotique » révèle un réseau d'allusions et d'échos lourds de sens. Proust emploie l'hyponyme « narcotique » lorsqu'il informe sa mère de l'état de santé de M. Cottin : « Il prend presque tous les soirs des narcotiques, étant très nerveux » (lettre_63_M). En parallèle, les deux occurrences de « cachet » sont associées à « trional » :

³⁰ M. Proust, *Jean Santeuil*, P. Clarac – Y. Sandre ed., Gallimard, Paris 1971 (Bibliothèque de la Pléiade, 228), pp. 205-206.

³¹ Pour le lexique du langage religieux dans l'œuvre de Proust : S. Chaudier, *Proust et le langage religieux. La cathédrale profane*, Honoré Champion, Paris 2004 (Recherches Proustiennes, 2).

³² Pour la comparaison dans la *Recherche* : I. Vidotto, *Proust et la comparaison vive. Étude stylistique*, Classiques Garnier, Paris 2020 (Bibliothèque proustienne, 28).

³³ M. Proust, *Les soixante-quinze feuillets*, pp. 32-33.

³⁴ Bien qu'elle n'appartienne pas à notre corpus épistolaire de référence, il incombe de mentionner ici une lettre de 1904 à la Princesse Alexandre de Caraman-Chimay, où l'image eucharistique est rapprochée du tétronal, un narcotique succédané du trional : « Quel présent mystérieux que ce tétronal. Par quelle communion incompréhensible la blanche hostie [...] m'apportera-t-elle pour quelques heures l'oubli des chagrins et me laissera-t-elle au matin, à l'heure du réveil, plus plein d'espérances et de résignation ? » (*Corr.*, XXI, p. 602-603).

Figure 3 - *Concordance* « cachet » Proust – Correspondance avec sa mère : 1887-1905

Tout comme *Jean Santeuil*, les « soixante-quinze feuillets » sont ancrés dans une perspective autobiographique. Il est possible que Proust pense au rapport avec sa mère lorsqu'il choisit d'accentuer la composante médicale du « baiser du soir ». La scène est développée ultérieurement – et d'une manière semblable – dans le Cahier 4 (1908) et dans le Cahier 8 (1909) :

une hostie où je goûterais sa présence réelle [...] ; douce hostie pour une communion de paix, qui m'assurerait pour toute la nuit un sommeil plus calme et plus doux que celui que nous trouvons dans ces autres hosties, bien miraculeuses aussi, que la pharmacie prépare [...] ³⁵.

Les deux isotopies de la religion et de la médecine sont largement modifiées. Certes, la « figure heureuse et tendre » de la mère est encore une fois comparée à la « nourriture » chrétienne par excellence : la « présence réelle » (doctrine qui évoque le mystère de la transsubstantiation) renferme la chair et le sang des « soixante-quinze feuillets ». Comme le sommeil induit par la mère est finalement « *plus* calme et *plus* doux », l'effet assoupissant des médicaments s'affaiblit : le « cachet de narcotique » disparaît du texte et les « autres hosties » se caractérisent plus génériquement par le fait d'être préparées par « la pharmacie » (ou par la « pharmacopée » ³⁶).

3.3. Le baiser du soir dans la *Recherche* : la disparition du bâton d'Esculape ?

L'éclipse progressive du domaine médical débouche sur la version du baiser du soir de la *Recherche* :

Or la voir fâchée détruisait tout le calme qu'elle m'avait apporté un instant avant, quand elle avait penché vers mon lit sa figure aimante, et me l'avait tendue comme une hostie pour une communion de paix où mes lèvres puiseraient sa présence réelle et le pouvoir de m'endormir ³⁷.

Les domaines religieux et médical, qui sont entrelacés dans la genèse, s'allient et puis se séparent : ils entrecroisent et mélangent leurs champs sémantiques pour suivre finalement deux chemins différents. De la « compresse » de *Jean Santeuil* aux « hosties préparées par la pharmacie » du Cahier 4, en passant par le « cachet de narcotique » des « soixante-quinze feuillets », le lexique médical disparaît de la version de la *Recherche*.

Cependant, le réservoir de connaissances que l'auteur mobilise dans sa correspondance et dans les versions antérieures du baiser du soir contamine d'autres lieux textuels. Par exemple, à l'instar du baiser de Mme Santeuil, la seule occurrence de

³⁵ Cahier 4, NAF 16644, 24r^o-25r^o, *Les soixante-quinze feuillets*, p. 159 (« Notes », p. 268).

³⁶ Cahier 8, NAF 16648, 16r^o ; M. Proust, *À la recherche du temps perdu*, vol. I, J.-Y. Tadié ed., Gallimard, Paris 1987-1989 (Bibliothèque de la Pléiade, 100), p. 680 (abrégé dorénavant par le sigle RTP suivi du tome en chiffre romain et du numéro de page en chiffre arabe).

³⁷ RTP, I, p. 13.

« compresse » dans la *Recherche* fait office de comparant : le souvenir de la vision d'Albertine devant Saint-Mars-le-Vêtu est « comme une grande compresse calmante qu'on eût appliquée à mon cœur »³⁸. Dans la *Prisonnière*, Bergotte « essaya avec succès, mais avec excès, de différents narcotiques », y compris ceux qui sont « dérivés [...] de l'amyle et de l'éthyle »³⁹. Justement, lorsqu'il passe un savon à M. de Cambremer, le docteur Cottard mentionne ces deux éléments à l'intérieur d'un long discours sur la posologie du trional (!) :

Vous parlez de trional, savez-vous seulement ce que c'est ? – Mais... j'ai entendu dire que c'était un médicament pour dormir. – Vous ne répondez pas à ma question [...]. Pouvez-vous me dire ce qu'il contient de parties d'amyle et d'éthyle ?⁴⁰.

Ce n'est pas un hasard que dans trois lettres sur les cinq qui mentionnent l'amyle, Proust convoque aussi le trional :

The screenshot shows a web-based concordance tool. At the top, it says 'CONCORDANCE' and 'Proust - Correspondance avec sa mère : 1887-1905'. Below that, there's a search bar with 'simple amyle • 5' and '70.9 per million tokens • 0.0071%'. The interface includes navigation icons and a table of results. The table has columns for 'Left context', 'KWIC', and 'Right context'. Five results are listed, each with a checkbox, a document ID, and a snippet of text where 'amyle' is highlighted in red.

	Left context	KWIC	Right context
1	doc#86 : de ton cabinet de toilette) poitrine gênée malgré fumage. 2 perles d'	amyle	. </s><s> Vite endormi. </s><s> Cinq heures 1/2 réveillé par oppressi
2	doc#88 ple bonjour. </s><s> Je continue à me passer rigoureuset de trional,	amyle	et valériane. </s><s> Mon dîner dans l'île avec Noufflard m'a si bien r
3	doc#89 </s><s> Après fumage (hélas j'avais déjà dû en dînant) je n'ai pris ni	amyle	, ni valériane, ni enfin rien (que tisane et bicarbonate) et je m'en suis :
4	doc#90 même onze h 1/4, couché à mi- nuit 20 très peu d'oppression, pris 2	amyle	préventivement et en effet nuit sans oppression du tout. </s><s> Mais
5	doc#91 ne à ce tt pett mot. </s><s> J'ai l'ennui de te dire que j'ai pris hier soir	amyle	et ce matin trional mais cela ne durera pas. </s><s> Je vais vraiment

Figure 4 - *Concordance* « amyle » Proust – Correspondance avec sa mère : 1887-1905

Enfin, la pharmacie de la *Recherche* offre une vaste gamme de cachets : « cachet de pyramidon » ; « cachets soporifiques » ; « cachet pour dormir » ; « cachets de véronal » et « cachets de caféine »⁴¹.

À la lumière du savoir médical dont le récit hérite, on peut affirmer que la scène du baiser du soir dans la *Recherche* est moins le résultat de l'abandon du domaine médical au profit du domaine religieux (qui est à son tour édulcoré et épuré des éléments plus blasphématoires) que l'aboutissement d'un long travail de reformulation et de refonte, d'intériorisation et de dissémination d'éléments médicaux dans le récit. Plus spécifiquement, la dissémination et l'occultation des références explicites au trional et aux narcotiques relèvent d'une démarche plus vaste d'occultation et de « cryptage »⁴² des données autobiographiques explicites, comme le prouve la confrontation avec la correspondance. C'est précisément parce que le trional fait partie du quotidien de Proust et de l'échange épistolaire avec sa mère que sa présence est gommée de la scène du baiser, et redistribuée ailleurs.

³⁸ RTP, III, p. 405.

³⁹ RTP, III, p. 691. On signale deux autres occurrences de « narcotique » : RTP II, p. 178 ; III, p. 373.

⁴⁰ RTP, III, p. 351.

⁴¹ RTP, II, p. 783 ; III, p. 121 ; p. 373 ; IV, p. 47.

⁴² F. Goujon, *Allusions littéraires et écriture cryptée dans l'œuvre de Proust*, Honoré Champion, Paris 2020 (Recherches proustiennes, 47).

Conclusion et perspectives

Cette étude a mis en exergue le potentiel d'une approche terminologique ayant comme but l'extraction d'unités lexicales spécialisées dans un corpus donné. Au vu des contenus des lettres que Marcel Proust échange avec sa mère, nous nous sommes focalisés sur le domaine médical. Le cas que nous avons étudié a montré non seulement la synergie entre la correspondance et le roman, mais aussi la façon dont ce dernier *adopte* les tecnicismes médicaux pour les *adapter* à la fois à la synchronie et à la diachronie du récit. Dans cette optique, l'édition numérique (*Corr-Proust*) permettra d'élargir le champ d'action à toute la correspondance proustienne, et de répondre donc à la question posée en ouverture de ce travail : « La médecine ne prouve-t-elle pas que la chronologie n'a rien à voir avec l'âge ? ». Dans la perspective des travaux à venir, nous nous proposons d'analyser d'autres termes simples (comme « fatigue ») ou complexes (comme « crise d'asthme ») pour interroger d'autres types de variation linguistique. Enfin, nous entendons mettre les résultats issus de cette étude à la disposition de la communauté scientifique. Dans la lignée des travaux menés précédemment sur la collecte et l'analyse de la terminologie médicale dans les œuvres de Sir Arthur Conan Doyle⁴³, nous visons à fournir une base de données terminologique, librement accessible en ligne, qui recueille le lexique médical spécialisé employé par Marcel Proust.

⁴³ F. Vezzani – G. M. Di Nunzio, *(Not so) Elementary, my dear Watson! A different perspective on medical terminology*, « Umanistica Digitale », 3(6), 2019 <https://doi.org/10.6092/issn.2532-8816/8632> (dernière consultation le 28 mars 2022).