



**UNIVERSITÉ
DE GENÈVE**

FACULTÉ DES LETTRES

UNIVERSITÉ DE GENÈVE
Faculté des Lettres
Département de linguistique

ANALYSE DES CHAMPS LEXICAUX DANS LE FONDS PATRIMONIAL DE JEAN-MARTIN CHARCOT

Mémoire présenté en vue de l'obtention du
Certificat de spécialisation en Linguistique

Étudiante :

Ljudmila PETKOVIC

N° de matricule : 19-337-757

ljudmila.petkovic@etu.unige.ch

Directeur :

Prof. Dr Christopher LAENZLINGER

Co-encadrant :

Dr Luka NERIMA

8 août 2024

TABLE DES MATIÈRES

Introduction	1
1 La rupture épistémologique en médecine : la notion d'hystérie	2
1.1 L'erreur comme source de la connaissance scientifique	2
1.2 Jean-Martin Charcot : un médecin polymathe à l'aube de la neurologie moderne	4
2 Pister la circulation du discours médical au prisme du numérique	7
2.1 Modalités des circulations des savoirs	7
2.2 Constitution d'un concept historique	8
2.3 Études numériques des circulations culturelles	9
3 Valorisation du fonds Charcot	11
3.1 Description du fonds Charcot	11
3.2 Constitution du corpus Charcot	13
4 Résultats	14
4.1 Exploration du corpus Charcot : OBVIE et TEXTPAIR	14
4.2 Extraction des phrases-clés : méthodes statistiques	14
4.3 Extraction des phrases-clés : méthode à base d'apprentissage profond . .	16
4.4 Discussion	16
Conclusion	17
Annexe	18
Liste des termes et expressions popularisées par Charcot	20
References	21

INTRODUCTION

motivation

Ce mémoire, à la jonction de l'histoire des sciences et de la linguistique computationnelle, propose une étude interdisciplinaire dont l'objectif est la valorisation numérique du fonds patrimonial de Jean-Martin Charcot, fondateur de la neurologie moderne au XIX^e siècle en France. Si l'importance de ses travaux scientifiques est un sujet largement étudié (Bogousslavsky, 2011 ; Broussolle *et al.*, 2012 ; Camargo *et al.*, 2024), il n'existe pas, à notre connaissance, d'initiatives en humanités numériques visant l'exploitation de son fonds volumineux et d'une grande importance scientifique. Nous souhaitons donc mesurer informatiquement l'impact scientifique des travaux de Charcot sur ses collaborateurs et successeurs, membres de son réseau scientifique¹. Cette mesure se fonde sur l'analyse des concepts-clés en matière de son discours scientifique, et plus particulièrement sur l'opérationnalisation du terme « influence », définie ici comme une intertextualité uni-directionnelle, allant des écrits de Charcot (ci-après corpus « Charcot ») vers ceux de ses collaborateurs et successeurs (ci-après corpus « Autres »). Il s'agit donc *in fine* d'aborder computationnellement la question des circulations, non pas des artefacts matériels comme les manuscrits (Gabay *et al.*, 2021) et les images (Joyeux-Prunel, 2019), mais des phénomènes textuels complexes (Manjavacas *et al.*, 2019) ayant une dimension théorique forte.

structure

Le présent mémoire est structuré en quatre parties principales : après l'introduction, nous traçons l'évolution du progrès médical dans le cadre épistémologique, où Charcot a joué un rôle important, avant de présenter ses contributions principales (chapitre 1). Dans le chapitre 2, nous soulignons les aspects des circulations des savoirs et proposons une revue de la littérature portant sur ce sujet du point de vue numérique. Le fonds Charcot et la constitution du corpus de recherche correspondant sont abordés dans le chapitre 3. Ensuite, le chapitre 4 présente les premières tentatives d'analyse computationnelle de l'impact de Charcot sur ses élèves et collègues, ainsi que les limites de ces approches, en proposant de nouvelles méthodes pour la quantification de la pertinence des expressions polylexicales. Enfin, le dernier chapitre est consacré à la conclusion du travail et aux pistes pour des recherches futures.

1. Par ailleurs, le présent travail fait partie du projet doctoral de l'autrice en cours, cf. <https://theses.fr/s382733>.

CHAPITRE 1 LA RUPTURE ÉPISTÉMOLOGIQUE EN MÉDECINE : LA NOTION D'HYSTÉRIE

1.1 L'erreur comme source de la connaissance scientifique

« Les vraies révolutions sont lentes et elles ne sont jamais sanglantes. »

— Anouilh (1956)

progrès de science

La science progresse en corrigeant constamment les erreurs, c'est-à-dire que les erreurs précèdent nécessairement l'établissement de la connaissance scientifique. Bien que ce processus de correction des erreurs puisse être observé de manière diachronique, il est de nature circulaire. En outre, si une doctrine devient obsolète avec le temps et l'avènement des technologies avancées permettant de recueillir de nouvelles preuves, une doctrine actuellement en vigueur deviendra tout de même à son tour obsolète à un moment¹.

changement de paradigme

Un tel cycle des observations empiriques peut être bouleversé, selon Bachelard (1934, p. 26), par la « rupture et non pas continuité entre l'observation et l'expérimentation ». Autrement dit, la rupture épistémologique survient lors d'un renversement fondamental dans la façon d'établir une connaissance dans un domaine particulier. De fait, ce phénomène caractérise une « révolution scientifique » (Koyré, 1957), terme apparenté avec celui du « changement de paradigme », introduit par Kuhn (1962). D'après ce dernier, les *paradigmes* désignent les « découvertes scientifiques universellement reconnues qui, pour un temps, fournissent à une communauté de chercheurs des problèmes types et des solutions ».

discontinuité

Dans cette optique, la structure des révolutions scientifiques désigne un modèle épistémologique constitué des épisodes non cumulatifs du développement scientifique (Figure

1. L'un des exemples le plus connu de l'obsolescence scientifique est sans doute le passage du modèle géocentrique de l'univers, défendu par Aristote et Ptolémée (selon lesquels la Terre est immobile au centre de l'Univers), à la conception héliocentrique de Nicolas Copernic, qui affirmait que la Terre tournait autour du Soleil.

1.1), marqués par des passages radicaux d'un paradigme à un autre. Le nouveau paradigme ne désigne donc pas une extension de l'ancien paradigme ; au contraire, ce dernier est entièrement ou partiellement remplacé par un nouveau paradigme incompatible avec le précédent. Cela est bel et bien un signe de l'émergence d'une nouvelle théorie ou découverte, tout en prouvant que le développement historique des théories est fondamentalement discontinu.

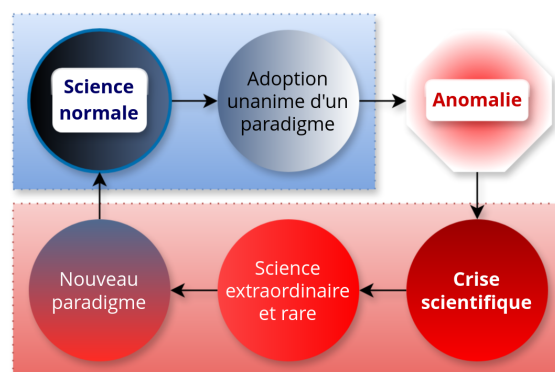


FIGURE 1.1 – Conception kuhnienne du progrès scientifique, adaptée de Amiri (2012).

Dans un esprit similaire, Bachelard (1970, p. 72), p. 72 souligne :

erreurs premières

« Il ne saurait y avoir de vérité *première*. Il n'y a que des erreurs *premières*. On ne doit donc pas hésiter à inscrire à l'actif du sujet son expérience essentiellement malheureuse. La première et la plus essentielle fonction de l'activité du sujet est de se tromper. Plus complexe sera son erreur, plus riche sera son expérience. L'expérience est très précisément le souvenir des erreurs rectifiées. L'être pur est l'être détrompé. »

évolution de l'hystérie

Un exemple du changement de paradigme est l'évolution du terme *hystérie*, introduit par Hippocrate dans l'Antiquité au V^e s. av. J.-C., qui expliquait cette maladie par un déplacement de l'utérus dans le corps féminin². Au Moyen Âge, surtout à partir du XIII^e s., les *hystériques* étaient considérées par l'Église comme possédées par le diable et, par conséquent, chassées, torturées ou soumises aux exorcismes dans une perspective religieuse (Tasca *et al.*, 2012). Néanmoins, certains scientifiques de la Renaissance commencent progressivement à s'éloigner de l'étiologie démonologique de cette maladie ; un cas notable est celui du médecin Charles Le Pois (1563-1633), qui fut le premier à désigner le cerveau, et plus précisément, le *sensorium commune*³, comme siège de la maladie hystérique en 1618⁴, en associant l'hystérie autant aux hommes qu'aux femmes (Wright, 1980).

2. Ce terme est issu du mot grec ὑστέρα, par le latin *hystera*, « matrice ». Par dérivation, le terme hystérique se référait à une personne « (femme) malade de l'utérus », selon Rey (2011).

3. ce que Kant (1863) appelle plus tard « siège commun de la sensibilité » pour désigner l'ensemble des perceptions.

4. Le Pois (1618, p. 101), p. 101 a noté que les symptômes communément appelés hystériques se référaient à l'épilepsie, mais qu'il était prouvé que l'épilepsie elle-même était une maladie *idiopathique* (existant par elle-même, sans lien avec une autre maladie) de la tête, et non pas provoquée par les troubles de l'utérus ou des intestins.

Pour mieux comprendre l'importance de ce changement de pensée radical, il convient également de souligner que notre compréhension actuelle du système nerveux central est basée sur les premières descriptions faites de manière rigoureuse par Constanzo Varolio (1543-1575) au XVI^e s. (Tubbs *et al.*, 2008)⁵. À l'époque des Lumières en Angleterre (fin XVII^e – début XVIII^e s.), Thomas Willis (1621-1675), créateur du terme *neurologia* en 1664⁶, maintint et développa cette conception en caractérisant cette maladie comme principalement convulsive en raison des explosions des « esprits animaux » dans le cerveau (Willis, 1681). Enfin, l'histoire de la neurologie trouve son ancrage à la fin du XIX^e siècle dans les travaux de Jean-Martin Charcot (1825-1893). Ce n'est qu'à cette période que la maladie en question a été systématiquement traitée comme un trouble neurologique (Tasca *et al.*, 2012). La sous-section 1.2 évoque certains de ses apports principaux dans le domaine scientifique et artistique.

trouble cérébral

1.2 Jean-Martin Charcot : un médecin polymathe à l'aube de la neurologie moderne

Figure emblématique et directeur de l'illustre École de la Salpêtrière (basée à l'actuelle hôpital de la Pitié-Salpêtrière à Paris), Charcot a laissé une trace indélébile dans le domaine de la neurologie. Il est essentiellement connu pour ses études portant sur les troubles névrotiques, notamment l'hystérie. Selon lui, l'hystérie découle d'une dégénérescence héréditaire du système nerveux (Tasca *et al.*, 2012), en montrant qu'elle est en fait plus fréquente chez les hommes que chez les femmes. Charcot a été reconnu pour ses travaux de recherche sur l'hypnose qu'il a utilisée afin d'induire l'état modifié de conscience d'un sujet, permettant l'analyse des symptômes hystériques, ainsi que comme méthode de traitement. Son nom est également associé aux descriptions de nombreuses pathologies connues aujourd'hui, comme la *maladie de Parkinson*, la *sclérose en plaques disséminées*, abbr. *SEP* (ou *sclérose multiple*), la *sclérose latérale amyotrophique*, abbr. *SLA* (soit la *maladie de Charcot*, ou *maladie Lou-Gehrig*) etc⁷.

contributions

Ces explorations des abîmes de l'esprit humain lui ont valu de nombreuses appellations : à part avoir été globalement considéré comme le père de la neurologie française et moderne (Teive *et al.*, 2022; Broussolle *et al.*, 2012), d'autres noms plus symboliques lui ont été associés, notamment « Napoléon des névroses », « Paganini de l'hystérie » (Marmion, 2015), ou même « César de la Faculté » (Camargo *et al.*, 2024). Dans la même lignée de pensée, l'École de la Salpêtrière était caractérisée comme la « Mecque de la neurologie » grâce aux activités de Charcot (Teive *et al.*, 2014; Goetz, 2017; Camargo *et al.*, 2024).

appellations

5. Il s'agit de l'identification et de la description de la structure cérébrale agissant comme un relai entre le cerveau et le cervelet, appelée *pont* (lat. *pons*) par Varolio (1573), soit *pont de Varole* (lat. *pons Varolii*), en l'honneur du célèbre anatomiste, qui fut le premier à examiner le cerveau de sa base vers le haut.

6. Cf. Willis (1664).

7. Pour un aperçu détaillé des contributions majeures de Charcot dans le domaine de la médecine, voir Camargo *et al.* (2024).

Charcot dans la littérature

En outre, de nombreuses références à Charcot et des descriptions d'attaques hystériques figurent non seulement dans la littérature médicale, mais aussi dans des romans naturalistes français et européens, notamment en Pays-Bas, Russie, pays scandinaves, Espagne, Italie et Allemagne (Koehler, 2013).

influences

Charcot a créé un véritable réseau scientifique et artistique autour de soi grâce à ses idées novatrices qui ont eu un grand retentissement parmi ses collaborateurs, élèves et savants polymathes, dont nous ne nommons que quelques figures majeures souvent citées dans la littérature (Gomes & Engelhardt, 2013 ; Bogousslavsky, 2014b ; Camargo *et al.*, 2024), notamment :

- Paul Richer (1849-1933), anatomiste, neurologue et sculpteur, qui a résumé les premières études de Charcot sur l'hystérie dans ses *Études cliniques sur l'hystéro-épilepsie ou grande hystérie* ;
- Georges Gilles de la Tourette (1857-1904), psychiatre et neurologue, qui a décrit les symptômes de la *maladie des tics*, renommée *syndrome de Tourette* en son hommage par Charcot lui-même ;
- Pierre Janet (1839-1916), philosophe, neurologue et psychiatre, concepteur des termes *dissociation* et *sous-conscient* ;
- Désiré Magloire Bourneville (1840-1909), homme politique et neurologue, qui a publié le premier tome de l'ouvrage monumental *l'Iconographie photographique de la Salpêtrière*, dédiée à l'hystérie, sous l'égide de Charcot ;
- Joseph Babinski (1857-1932), neurologue et neurobiologiste, concepteur du terme *pithiatisme*, qui a découvert le réflexe cutané plantaire, appelé également *signe de Babinski*.

avis partagés

L'impact colossal de Charcot sur sa propre discipline se reflète aussi dans le changement d'intérêt radical du célèbre psychanalyste Sigmund Freud (1856-1939), caractérisé par le passage de la neurologie générale à l'hystérie, l'hypnose et d'autres troubles psychologiques. En effet, son séjour dans le service de Charcot à Paris en 1885-1886 a donné lieu au développement de la théorie psychanalytique (Camargo *et al.*, 2018). Néanmoins, certains scientifiques ont fortement contesté le raisonnement scientifique de Charcot, comme le neurologue Hippolyte Bernheim (1840-1919) avec l'École de Nancy pendant les années 1880-1890. Cette polémique porte sur la nature de l'hypnose qui, pour Charcot, représentait un état pathologique propre aux hystériques, et non pas un état de sommeil obtenu par suggestion qui est susceptible d'applications thérapeutiques (et donc, applicable à pratiquement n'importe qui), comme le soutenait Bernheim (1891).

pont

Étant donné l'interdisciplinaire des travaux de Charcot et ses contributions dans le domaine de la neurologie, nous souhaitons explorer la notion de la circulation des savoirs au prisme du numérique à travers son impact. Avant d'aborder la question d'opérationnalisation de son impact, nous tenons d'abord à décortiquer les mécanismes à l'origine des circulations des savoirs à grande échelle, ainsi que de définir la notion d'un « concept » pouvant véhiculer les informations importantes concernant les circulations

en question.

CHAPITRE 2 PISTER LA CIRCULATION DU DISCOURS MÉDICAL AU PRISME DU NUMÉRIQUE

2.1 Modalités des circulations des savoirs

De nombreux·ses chercheur·se·s·x partagent le point de vue selon lequel la notion de « circulation des savoirs » constitue un champ de recherche vaste, ainsi qu'un nouveau paradigme de la connaissance depuis le début du XXI^e siècle et l'avènement du Web 2.0¹ (Landais, 2014; Quet, 2014). Le terme en question reste toutefois assez complexe en raison de visions différentes sur la façon de le définir. Afin d'éclairer cette problématique,

3 aspects

Quet (2014) souligne trois aspects suivants :

1. **Éléments de la circulation.** Qu'est-ce qui circule ?
 - individus (savants, techniciens, traducteurs, etc.)
 - objets matériels (instruments scientifiques, ouvrages etc.)
 - constructions symboliques (théories, concepts etc.)
2. **Conceptions de la circulation et méthodes de son analyse**
 - définition de la circulation comme « traduction », « diffusion », « accès » ou « succès »
 - critères méthodologiques possibles pour étudier la circulation p. ex. d'une théorie :
 - circulations géographiques des principaux concepteurs qu'on lui reconnaît
 - circulations et lectures des textes produits par leurs concepteurs
 - usages et applications analogiques qui en sont faits dans d'autres domaines
 - enjeux d'articulation de ces différents niveaux d'observation du point de vue méthodologique et de celui de la production du texte de recherche, dans le cas des croisements de ces niveaux
3. **Conceptions analytiques et normatives des savoirs**
 - affaiblissement des catégories des « savoirs profanes » et « savoirs scientifiques », ainsi que de l'opposition entre eux

1. Cette phase de l'évolution du Web se caractérise notamment par la transformation majeure de l'Internet en vue du développement des réseaux sociaux, des blogs et des sites participatifs, tout en permettant aux utilisateur·trice·s·x de créer, partager et interagir avec du contenu Web. Nous traversons actuellement l'ère du Web 3.0 qui repose sur des technologies telles que la chaîne de blocs (angl. *blockchain*), le NFT (angl. *non-fungible token*), l'intelligence artificielle, métavers et le Web sémantique (Varet, 2023).

- revalorisation des savoirs implicites et de la dimension pratique des connaissances
- glorification de la circulation comme porteuse de valeurs *a priori* positives : confrontation à l'autre, hybridation, production de nouveauté, etc.

Dans le cadre de l'analyse de l'impact scientifique de Charcot, nous étudions *in fine* la circulation de ses théories et des concepts médicaux dont il était inventeur (p. ex. *SLA*) et transmetteur (p. ex. *hystérie*)². La section 2.2 élabore les différentes approches pour définir plus globalement la notion des concepts historiques qui nous orienteront vers une définition des concepts médicaux en particulier.

pont

2.2 Constitution d'un concept historique

Le mot « concept » est un terme générique qui renvoie à un grand nombre de définitions provenant de divers domaines de pensée, sans qu'il en existe une qui soit exhaustive et universellement acceptée. En mathématiques et en sciences cognitives (philosophie, psychologie, intelligence artificielle etc.), un concept représente une catégorie basique axiomatique qui est indétectable et comprise de manière intuitive. On peut le comprendre également comme une unité de connaissance, une idée, un cadre, un script, un *gestalt*³ ou un hypéronyme d'une notion (Лихачёв, 1997).

concept mathématique

En revanche, selon les linguistes, un concept a une structure double, constituée du sens linguistique et culturel (Nemickienė, 2011). Sa couche intérieure est constituée du noyau étymologique sur lequel repose ensuite la couche périphérique qui hérite les éléments formés par la culture, les traditions et les expériences humaines (Степанов, 2007). En linguoculturologie, on retrouve le terme « concept linguo-culturel » qui reflète cette nature double du concept. Il peut être exprimé par de différents éléments du langage, soit : lexèmes, idiomes, collocations, phrases ou textes entiers.

concept linguistique

Concept / circulation en analyse du discours / TAL / HN?

- TAL : concept = groupe de mots, créé manuellement ou extrait à partir des ontologies ou des thésaurus

Afin de pouvoir analyser les concepts médicaux liés à Charcot, il est essentiel de déterminer à partir de quel moment un mot ou un groupe de mots devient un concept en sciences humaines et sociales (ci-après SHS). Du point de vue de l'histoire des concepts (allem. *Begriffsgeschichte*), cette transformation survient lorsqu'un seul mot comprend toute la gamme des significations dérivées d'un contexte sociopolitique (Koselleck & Richter, 2011, p. 258). À titre d'exemple, le concept d'un *état* ne peut être interprété qu'à travers ses différents constituants, dont *souveraineté territoriale, législation, fiscalité*, parmi

Begriffsgeschichte

2. Comme déjà expliqué dans la sous-section 1.1, Charcot n'a pas inventé ce terme, mais en réinterprété le sens.

3. Allem. « forme, figure ». Le terme provient du domaine de la psychologie de la forme (théorie de la *gestalt* ou *gestaltisme*) dans les années 1920. Selon ce courant, un individu observe les phénomènes dans leur ensemble de manière holistique, et non pas comme les parties individuelles juxtaposées ou additionnées (p. ex. cinq cercles qui se chevauchent, lesquels l'œil humain perçoit comme le logo se référant au concept des Jeux olympiques).

maints d'autres. Les concepts sont donc les concentrations par défaut ambiguës d'une multitude de contenus sémantiques, uniquement interprétables et indéfinissables, par contraste avec des significations des mots qui peuvent être définies de manière exacte (Koselleck & Richter, 2011, p. 20).

De plus, les concepts comme *histoire* ou *progrès* sont caractérisés comme « collectifs singuliers » qui marquent un passage du domain concret d'un individu (plusieurs *histoires* et *progrès* individuels) au domain abstrait et général du collectif social (une *histoire* ou un *progrès* général ou collectif). Ce phénomène linguistique, ainsi que la création des concepts comme *industrie*, *usine*, *classe moyenne* etc., reflète un changement de paradigme dans l'organisation sociale survenu lors des révolutions politiques et industrielles (Hobsbawm, 2010, p. 1). Cela traduit donc le lien fort entre l'histoire du langage et l'histoire des idées. **La conceptosphere est un ensemble des concepts considérés comme caractéristique d'une nation particulière.** Koselleck nomme cette période charnière *Sattelzeit*⁴ (2011, p. 8, p. 8), entre 1750 et 1830, durant laquelle les concepts historiques deviennent abstraits, singularisés, respatialisés et retemporalisés.

Ces considérations peuvent s'appliquer à d'autres constructs en SHS, comme *travail*, *intelligencija*, *Ancien Régime*, *avant-garde*, *Occident* etc. Elles ont acquis le statut des concepts « nomades » en raison de leur circulation spatio-temporelle et linguistique (Ghermani, 2011, p. 117). Plusieurs questionnements ont été soulevés par la même autrice à l'égard de leur émergence, notamment pour déterminer à quel moment un concept devient une entrée dans un dictionnaire des SHS : « Pourquoi un concept fait-il son entrée dans un dictionnaire ? Au terme de quel processus ? À l'inverse, comment cette percée lexicale est-elle parfois impossible ou refusée ? ». Les processus permettant à un concept d'obtenir le statut de scientificité sont la propagation, la bifurcation, la capture⁵, mais aussi les pratiques scientifiques conduisant aux masquages de sens (p. ex. dans le cas du terme « confession [religieuse] », dont le sens varie en fonction du pays dans lequel il est utilisé).

2.3 Études numériques des circulations culturelles

Incontestablement, l'époque actuelle est profondément marquée par le « déluge des données », phénomène représentatif de la quatrième paradigme de la science, selon Jim Gray (Hey *et al.*, 2009). Par conséquent, les recherches numériques sont aujourd'hui « pilotées par les données »⁶ et celles qui sont centrées sur les circulations culturelles se concrétisent à grande échelle.

Les humanités numériques au service de l'analyse des circulations culturelles se manifestent sous forme de divers projets de recherche au niveau académique. Certains éta-

4. Trad. allem. « époque de selle ».

5. Termes employés par Stengers (1987), représentatrice de la conception constructiviste du savoir scientifique.

6. Traduction du terme « data-driven » introduit par (Johns, 1991) issu du terme *data-driven learning*.

blissements universitaires, comme la chaire des Humanités numériques à l'université de Genève (Joyeux-Prunel & Gabay, 2022), ainsi que différents événements scientifiques (Humanistica 2023⁷, ACFAS 2023⁸ etc.) sont fortement axés sur cette thématique.

7. <https://humanistica2023.sciencesconf.org/>

8. <https://www.crihn.org/nouvelles/2022/12/11/colloque-de-la-transformation-des-sciences-humaines-par-les-humanites-numeriques-acfas-2023/>

CHAPITRE 3 VALORISATION DU FONDS CHARCOT

3.1 Description du fonds Charcot

fonds Charcot

Le fonds patrimonial de Jean-Martin Charcot est conservé à la Bibliothèque de Neurosciences Jean-Martin Charcot par la Bibliothèque numérique patrimoniale de Sorbonne Université (BSU)¹. Ce fonds regroupe des ouvrages suivants :

- fonds historique Charcot (bibliothèque personnelle de Charcot) : ouvrages, périodiques, collection de thèses et de tirés à part, manuscrits, observations, collection neurologique couvrant la seconde partie du XIX^e siècle, fonds bibliophilique ancien ;
- collections de la bibliothèque des Internes de la Salpêtrière : ouvrages, périodiques, thèses en neurologie et psychiatrie pour la période 1800-1950 ;
- donations en ouvrages du docteur Achille Souques.

Dans un souci de préservation d'ouvrages originaux et de valorisation de collections ayant un caractère iconographique notable, une partie de ce fonds a été numérisée. Ces archives numérisées sont disponibles sur le portail numérique SorbonNum², porte d'entrée unique vers les collections scientifiques patrimoniales et numériques de Sorbonne Université, ainsi que sur Gallica, bibliothèque numérique de la Bibliothèque nationale de France (BNF)³.

SorbonNum

Gallica

4 typologies de docs

Le fonds numérisé a été décrit et divisé par la BSU en quatre grandes typologies de documents :

1. Fonds iconographique

- **Album des internes** : Album des promotions annuelles d'internes, photographiées et classées par établissements de l'Assistance Publique, entre 1860 et 1963 ;
- **Photographies sur les aliénés de Bicêtre par Désiré Magloire Bourneville** :

1. <https://www.sorbonne-universite.fr/bu/decouvrir-nos-bibliotheques/la-bibliotheque-charcot>.

2. anc. Bibliothèque numérique patrimoniale de l'université Pierre et Marie Curie, <https://patrimoine.sorbonne-universite.fr/collection/Fonds-Charcot>

3. <https://gallica.bnf.fr/services/engine/search/sru?operation=searchRetrieve&version=1.2&query=%28gallica%20all%20%22Charcot%2C%20Jean-Martin%22%29&lang=fr&suggest=0>.

deux albums présentant les photographies des « petits enfants anormaux » hospitalisés à Bicêtre dans le service du docteur Bourneville, collaborateur de Charcot.

2. Leçons et manuscrits des leçons de Charcot

- **Manuscrits des leçons et observations de Charcot (1825-1893)** : leçons orales de Charcot, rédigées intégralement de sa main et annotées ;
- **Leçons de Charcot** : numérisation des volumes de l'*Œuvre Complète* de Charcot consacrés au système nerveux et à l'enseignement clinique, comme par exemple les célèbres leçons du Mardi, sur l'hystérie notamment.

3. Périodiques

- ***Les Recherches cliniques et thérapeutiques sur l'épilepsie, l'hystérie et l'idiotie (1872-1903)*** de Bourneville. Y est retracée toute l'activité du Service des Enfants Idiots, à la Salpêtrière puis à Bicêtre, par le biais des compte-rendu illustrés de photographies et rédigés par Bourneville ;
- ***Revue de l'Hypnotisme (1887-1910)*** : périodique consacré à l'hypnotisme que Charcot a réhabilité, publiant les principaux articles théoriques sur cette discipline ;
- ***Journal du magnétisme (1845-1861)*** : la collection reflète les recherches sur le magnétisme, renouvelées au milieu du XIX^e siècle ;
- ***Revue photographique des hôpitaux de Paris (1869-1872)***. Première revue exposant les applications de la photographie à la médecine, notamment la médecine hospitalière, à travers les études menées à l'Hôpital Saint Louis, et à la Salpêtrière ;
- ***Iconographie Photographique de la Salpêtrière (1875-1879)***. La collection présente les observations de patientes examinées à la Salpêtrière, accompagnées de photographies d'Albert Londe, présentant les divers stades de la crise d'hystérie ;
- ***Nouvelle Iconographie de la Salpêtrière (1888-1918)***. La revue est fondée sous la direction de Charcot par Paul Richer, Gilles de la Tourette et Albert Londe, directeur du service photographique. Elle réunit la collection de clichés constituée à la Salpêtrière a pour but la représentation objective des pathologies observées. Elle prend la relève de l'*Iconographie Photographique de la Salpêtrière*. Les articles sont illustrés de photographies, de dessins et de lithographies ;
- ***Archives de neurologie (1880-1907)***. Sous-titrée « Revue trimestrielle des maladies nerveuses et mentales », les Archives de neurologie sont publiées sous la direction de Charcot par Bourneville. La revue édite, groupe, catégorise et compare la masse des travaux de pathologie nerveuse. Les *Archives de neurologie* sont devenues bisannuelles en 1881.

4. Ouvrages de la bibliothèque de Charcot

- **Collection d'atlas d'anatomie et de pathologie du système nerveux**, publiés durant le XIX^e siècle. L'iconographie de ces ouvrages est remarquable, à commencer par l'*Atlas de Vicq d'Azyr*, médecin du roi Louis XVI ;
- **Traités**. Cette collection regroupe à la fois des traités sélectionnés dans la bibliothèque de Charcot (comme l'*Opera omnia*. . . de Thomas Willis, 1682, comportant

des gravures), des atlas et des textes significatifs des successeurs de Charcot, issus de la bibliothèque des Internes de la Salpêtrière (par exemple l'*Anatomie des centres nerveux* des Déjerine).

3.2 Constitution du corpus Charcot

Le corpus de travail est constitué de 201 documents OCRisés (sans post-correction), fournis au format XML par la BSU. Nous avons procédé, dans un premier temps, à une restructuration des textes en XML-TEI à l'aide de l'outil TEINTE⁴, afin de permettre la fouille avancée du corpus Charcot à travers des outils développés au sein de l'équipe-projet OBTIC. D'une part, le moteur de recherche OBVIE⁵ permet de repérer des textes similaires par ordre de pertinence à partir des termes en commun. D'autre part, l'algorithme TEXTPAIR génère une liste de passages similaires, c'est-à-dire les séquences de mots qui se chevauchent (n-grammes de mots) pour chaque texte, en comparant ensuite ces résultats avec ceux de séquences dans d'autres textes⁶.

Afin de mesurer l'impact de Charcot sur son entourage et d'analyser la circulation de concepts véhiculés dans le corpus, nous avons commencé par séparer les documents rédigés par Charcot de ceux rédigés par ses co-auteurs (p. ex. Bourneville) ou les auteurs thématiquement proches de lui (p. ex. de la Tourette). Nous avons obtenu respectivement 68 (corpus « Charcot ») et 133 (corpus « Autres ») documents, comme présenté dans le tableau 3.1.

Corpus	Nombre de documents	Nombre de tokens
Charcot textes rédigés par Charcot	68	12 190 649 (38,12%)
Autres textes rédigés par les membres de son réseau scientifique	133	19 788 830 (61,88%)
TOTAL	201	31 979 479 (100%)

TABLEAU 3.1 – Répartition du fonds Charcot selon les auteurs.

Les deux corpus issus du fonds Charcot sont librement disponibles et interrogeables sur les deux plateformes OBVIE⁷ et TEXTPAIR⁸.

4. https://github.com/OBVIL/teinte_obtic

5. <https://obtic.huma-num.fr/obvie/>. Pour d'amples informations sur le fonctionnement de cet outil, cf. Alrahabi (2022).

6. <https://artfl-project.uchicago.edu/text-pair>.

7. <https://obtic.huma-num.fr/obvie/charcot/?view=corpus>

8. <https://anomander.uchicago.edu/>

CHAPITRE 4 RÉSULTATS

4.1 Exploration du corpus Charcot : OBVIE et TEXTPAIR

Une première exploration du corpus Charcot à travers l'application OBVIE nous a permis d'identifier les substantifs les plus importants de chaque corpus en utilisant les fréquences brutes ou des méthodes plus fines comme TF-IDF, BM25 (détaillées dans la partie 4.2), χ^2 ou le TEST GAMMA. Cependant, l'application ne permet pas de quantifier la pertinence des expressions polylexicales, soit les n-grammes de mots, très fréquentes dans les deux corpus et dont la décomposition entraînerait une perte d'information (p. ex. le terme polysémique « bulbe » qui a une valeur spécifique dans l'expression figée *bulbe rachidien*). En observant la figure 4.1, nous constatons que l'abscisse donne l'information sur les dates de publication des ouvrages compris dans les corpus, alors que l'ordonnée indique le nombre d'occurrences par million de mots, soit *parties par million* (ppm)¹.

Concernant l'alignement des séquences similaires aux deux corpus, TEXTPAIR nous a permis, par une lecture attentive, de faire des comparaisons entre les textes et de rechercher des termes au sein des passages similaires, malgré le nombre de résultats assez conséquent (cf. la figure 4.2).

4.2 Extraction des phrases-clés : méthodes statistiques

Afin de surmonter les limites rencontrées avec ces deux outils, nous avons proposé une nouvelle méthode pour identifier des concepts dans les deux corpus en nous basant sur le poids de leur apparition, calculé selon trois différentes mesures de pondération² :

- TF-IDF est une méthode qui permet d'évaluer l'importance d'un terme contenu dans un document relativement à un corpus plus large en récompensant la fréquence des termes, sans tenir compte des variations de longueur du document;

1. Cf. le guide d'utilisation d'OBVIE détaillé <https://obtic.huma-num.fr/obvie//static/aide.html>.

2. Le code est disponible en ligne : https://github.com/ljpetkovic/Charcot_circulations.



FIGURE 4.1 – Distribution des fréquences des tokens avec la frise chronologique pour ceux constituant l’expression *bulbe rachidien* (issus du corpus « Charcot » et du corpus « Autres ») dans le logiciel OBVIE.

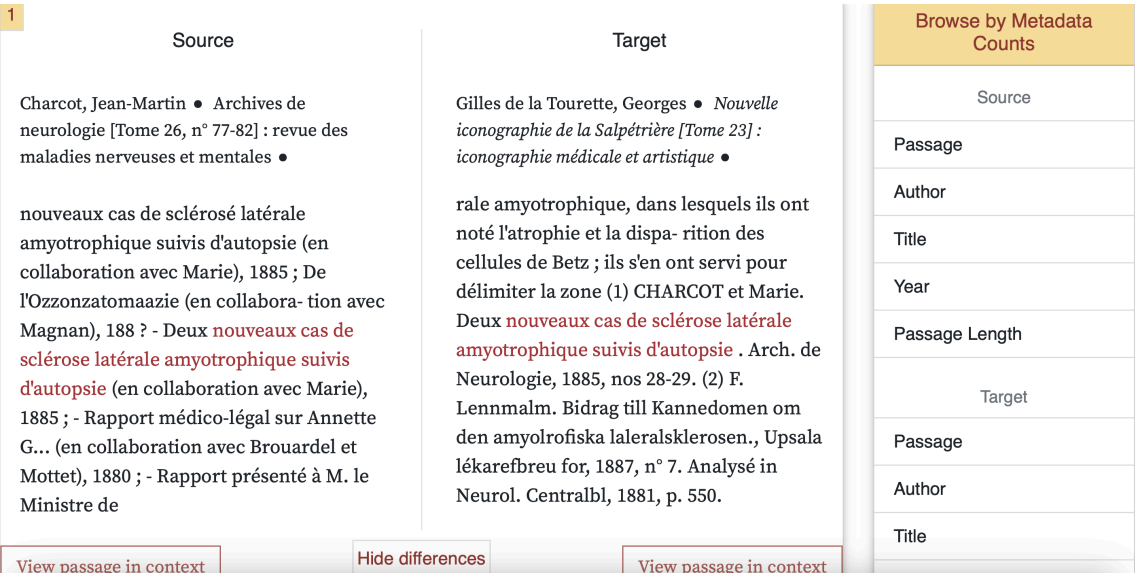


FIGURE 4.2 – Alignement et comparaison des textes de Charcot à celui de Georges Gilles de la Tourette (le seul résultat) en lançant la requête *sclérose latérale amyotrophique*.

- BM25 est une fonction de classement qui classe un ensemble de documents en fonction des termes de requête apparaissant dans chaque document, quelle que soit l’interrelation entre les termes de requête au sein d’un document (par exemple, leur proximité relative). Il s’agit d’une tentative d’amélioration de TF-IDF, notamment pour prendre en compte divers facteurs tels que la longueur du document et les problèmes engendrés par la possible saturation des termes (Robertson & Jones, 1976);
- BERT (Devlin *et al.*, 2019) est un modèle pré-entraîné qui utilise l’apprentissage non-supervisé sur de grandes quantités de données textuelles pour apprendre des représentations de mots et de phrases, et comprendre le contexte et la sémantique. Il

est basé sur l'architecture des *transformeurs*, qui est un type de réseau de neurones utilisé pour le traitement du langage naturel.

4.3 Extraction des phrases-clés : méthode à base d'apprentissage profond

4.4 Discussion

CONCLUSION

Intro / Rappel Contexte

Nous avons donc pu en tirer la problématique suivante :

CONCLUSION

Annexe

ANNEXE

Liste des termes et expressions popularisées par Charcot

BIBLIOGRAPHIE

- Alrahabi, M. (2022). Obvie : interface web pour la fouille et la comparaison de textes. In *Atelier Digital Humanities and cultural heritage : data and knowledge management and analysis during the francophone conference on the Extraction and the Management of Knowledge (egc2022)*. <https://hal.science/hal-03543362/>. (page 13)
- Amiri, V. V. (24 novembre 2012). T. S. Kuhn. *Histo Philo Sciences*. <https://histoirephilosciences.wordpress.com/depuis-le-20eme-siecles/une-nouvelle-epistemologie/t-s-kuhn/>. (page 3)
- Anouilh, J. (1956). *Pauvre Bitos ou le dîner de têtes*. Gallimard, coll. « Folio », n° 301. <https://archive.org/details/anouilh-pauvre-bitos-ou-le-diner-de-tetes-1979>. (page 2)
- Bachelard, G. (1934). *La formation de l'esprit scientifique : contribution à une psychanalyse de la connaissance*. Vrin. https://gastonbachelard.org/wp-content/uploads/2015/07/formation_esprit.pdf. (page 2)
- Bachelard, G. (1970). *Idéalisme discursif*. Vrin, présentation de Georges Canguilhem : Paris. https://www.academia.edu/27217437/BACHELARD_Gaston_%C3%89tudes_Vrin_1970_. (page 3)
- Bernheim, H. (1891). *De la suggestion et de ses applications à la thérapeutique*. Paris : Octave Doin. <https://gallica.bnf.fr/ark:/12148/bpt6k97805169>. (page 5)
- Bogousslavsky, J. (2011). *Following Charcot : A Forgotten History of Neurology and Psychiatry*, volume 29. Karger Medical and Scientific Publishers. <https://nah.sen.es/en/issues/latest-issues/135-journals/volume-2/issue-2/270-the-mysteries-of-hysteria>. (page 1)
- Bogousslavsky, J. (2014a). Jean-Martin Charcot and His Legacy. In *Hysteria : The rise of an Enigma*, volume 35 (pp. 44–55). Karger Publishers. <https://doi.org/10.1159/000359991>.

- Bogousslavsky, J. (2014b). The Mysteries of Hysteria. *Neurosciences and History*, 2(2), 54–73. https://nah.sen.es/vmfiles/abstract/NAHV2N2201454_73EN.pdf. (page 5)
- Broussolle, E., Poirier, J., Clarac, F., & Barbara, J.-G. (2012). Figures and institutions of the neurological sciences in Paris from 1800 to 1950. Part III : Neurology. *Revue Neurologique*, 168(4), 301–320. <https://doi.org/10.1016/j.neurol.2011.10.006>. (pages 1, 4)
- Camargo, C. H. F., Coutinho, L., Correa Neto, Y., Engelhardt, E., Maranhão Filho, P., Walusinski, O., & Teive, H. A. G. (2024). Jean-Martin Charcot : the polymath. *Arquivos de Neuro-psiquiatria*, 81, 1098–1111. <https://www.thieme-connect.de/products/ejournals/pdf/10.1055/s-0043-1775984.pdf>. (pages 1, 4, and 5)
- Camargo, C. H. F., Marques, P. T., de Oliveira, L. P., Germinian, F. M., de Paola, L., & Teive, H. A. G. (2018). Jean-Martin Charcot's Influence on Career of Sigmund Freud, and the Influence of this Meeting for the Brazilian Medicine. *Revista Brasileira de Neurologia*, 54(2). <https://docs.bvsalud.org/biblioref/2018/07/907032/revista542v4-artigo6.pdf>. (page 5)
- Christin, O. (2011). *Dictionnaire des concepts nomades en sciences humaines*. Mé-tailié. https://www.academia.edu/31022748/Dictionnaire_des_concepts_nomades_en_sciences_humaines_vol_1.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). Bert : Pre-training of Deep Bidirectional Transformers for Language Understanding. (page 15)
- Лихачёв, Д. С. (1997). Концептосфера русского языка. *Русская словесность: Ан-тология*. https://www.lihachev.ru/pic/site/files/fulltext/ocherk_po_philos/13.pdf. (page 8)
- Степанов, Ю. С. (2007). *Концепты. Тонкая пленка цивилизации*. Москва: Языки сла-вянских культур. (page 8)
- Gabay, S., Petkovic, L., Bartz, A., Levenson, M. G., & Du Noyer, L. R. (2021). Katabase : À la recherche des manuscrits vendus. In *Humanistica 2021* (pp. 1–7). <https://hal.science/hal-03066108>. (page 1)
- Ghermani, N. (2011). Confessions. In O. Christin (Ed.), *Dictionnaire des concepts nomades en sciences humaines* (pp. 117–133). Mé-tailié. https://www.academia.edu/5335160/_Confession_. (page 9)
- Goetz, C. (2017). Charcot : Past and present. *Revue Neurologique*, 173(10), 628–636. <https://doi.org/10.1016/j.neurol.2017.04.004>. (page 4)
-

- Gomes, M. d. M. & Engelhardt, E. (2013). Jean-Martin Charcot, father of modern neurology : an homage 120 years after his death. *Arquivos de Neuro-Psiquiatria*, 71, 815–817. <https://doi.org/10.1590/0004-282X20130128>. (page 5)
- Hey, T., Tansley, S., & Tolle, K. M. (2009). Jim Gray on eScience : A Transformed Scientific Method. In T. Hey, S. Tansley, & K. M. Tolle (Eds.), *The Fourth Paradigm*. Microsoft Research. <http://languagelog.ldc.upenn.edu/myl/JimGrayOnE-Science.pdf>. (page 9)
- Hobsbawm, E. (2010). *Age of revolution : 1789-1848*. Hachette UK. <https://files.libcom.org/files/Eric%20Hobsbawm%20-%20Age%20of%20Revolution%201789%20-1848.pdf>. (page 9)
- Johns, T. F. (1991). Should you be persuaded. two samples of data-driven learning materials. <https://api.semanticscholar.org/CorpusID:53988458>. (page 9)
- Joyeux-Prunel, B. (2019). Visual Contagions, the Art Historian, and the Digital Strategies to Work on Them. *Artl@s Bulletin*, 8(3), 128–144. <https://docs.lib.purdue.edu/artlas/vol8/iss3/8/>. (page 1)
- Joyeux-Prunel, B. & Gabay, S. (2022). Circulations des savoirs, de la recherche à l'enseignement. *Arabesques*. <https://doi.org/10.35562/arabesques.2847>. (page 10)
- Kant, É. (1863). *Anthropologie d'un point de vue pragmatique* (trad. J. Tissot). Librairie Ladrangé (originellement publié en 1798). https://fr.wikisource.org/wiki/Anthropologie_d%E2%80%99un_point_de_vue_pragmatique. (page 3)
- Kneib, M. (2011). *Étude fonctionnelle d'un circuit inhibiteur du cortex cérébelleux de la souris : Importance pour la physiopathologie des retards mentaux*. PhD thesis, Strasbourg. https://publication-theses.unistra.fr/public/theses_doctorat/2011/KNEIB_Marie_2011.pdf.
- Koehler, P. J. (2013). Chapter 6 – Charcot, La Salpêtrière, and Hysteria as Represented in European Literature. In S. Finger, F. Boller, & A. Stiles (Eds.), *Literature, Neurology, and Neuroscience : Neurological and Psychiatric Disorders*, volume 206 of *Progress in Brain Research* (pp. 93–122). Elsevier. <https://www.sciencedirect.com/science/article/pii/B9780444633644000235>. (page 5)
- Koselleck, R. & Richter, M. (2011). Introduction and Prefaces to the *Geschichtliche Grundbegriffe* : (Basic Concepts in History : A Historical Dictionary of Political and Social Language in Germany). *Contributions to the History of Concepts*, 6(1), 1–37. <https://www.berghahnjournals.com/view/journals/contributions/6/1/choc060102.xml>. (pages 8, 9)
-

- Koyré, A. (1957). *From the Closed World to the Infinite Universe*, volume 1. Baltimore, Johns Hopkins Press. <https://archive.org/details/fromclosedworldt0000koyr/page/n13/mode/2up?q=revolution>. (page 2)
- Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press. <https://www.lri.fr/~mbl/Stanford/CS477/papers/Kuhn-SSR-2ndEd.pdf>. (page 2)
- Landais, É. (2014). « Frédéric Darbellay, éd., *La circulation des savoirs. Interdisciplinarité, concepts nomades, analogies, métaphores* » : Berne, P. Lang, 2012, 245 pages. *Questions de communication*, 26, 331–333. <https://doi.org/10.4000/questionsdecommunication.9367>. (page 7)
- Le Pois, C. (1618). *Selectiorum observationum et consiliorum de praetervisis hactenus morbis affectibusque praeter naturum, ab aqua seu serosa colluvie et diluvie ortis, liber singularis*. Authore Carolo Pisone, Ponte ad Monticulum, apud Carolum Mercatorem. https://archive.org/details/BIUSante_05814/page/n3/mode/2up. (page 3)
- Manjavacas, E., Long, B., & Kestemont, M. (2019). On the Feasibility of Automated Detection of Allusive Text Reuse. In *Proceedings of the 3rd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature* (pp. 104–114). Minneapolis, USA : Association for Computational Linguistics. <https://doi.org/10.18653/v1/W19-2514>. (page 1)
- Marmion, J.-F. (2015). Freud hypnotiseur. In *Freud et la psychanalyse* (pp. 22–29). : Sciences Humaines. <https://www.cairn.info/freud-et-la-psychanalyse-9782361063542-page-22.htm>. (page 4)
- Nemickienė, Ž. (2011). “Concept” in Modern Linguistics : the Component of the Concept “Good”. *Filologija*, 16, 26–36. <https://core.ac.uk/outputs/62656539?source=oai>. (page 8)
- Quet, M. (2014). « Frédéric Darbellay, *La circulation des savoirs. Interdisciplinarité, concepts nomades, analogies, métaphores* ». *Revue d’anthropologie des connaissances*, 8(8-1). <https://doi.org/10.3917/rac.022.0221>. (page 7)
- Rey, A. (2011). *Dictionnaire historique de la langue française*. Le Robert. <https://ia601001.us.archive.org/2/items/alainreyetal.dictionnairehistoriquedelalanguefrancaise4eed.lerobert2010/Alain%20Rey%20et%20al.%20-%20Dictionnaire%20historique%20de%20la%20langue%20francaise%204e%20%C3%A9d.%20-%20Le%20Robert%20%282010%29.pdf>. (page 3)
- Robertson, S. E. & Jones, K. S. (1976). Relevance Weighting of Search Terms. *Journal of the American Society for Information science*, 27(3),

- 129–146. https://asistdl.onlinelibrary.wiley.com/doi/pdf/10.1002/asi.4630270302?casa_token=TfyVkMGkDQsAAAAA:TCuXWzGHjo31RdxGR9jECRG2rZzqv0K3G0zHF7yAa2NfxtDFqxe-MmSHMC6e80FiFxI4sLj2aW60yDk. (page 15)
- Roudinesco, É. & Plon, M. (2023). *Dictionnaire de la psychanalyse*. Fayard. <https://www.fayard.fr/sciences-humaines/dictionnaire-de-la-psychanalyse-nouvelle-edition-9782213725277>.
- Stengers, I. (1987). D'une science à l'autre. : Des concepts nomades. <https://archive.org/details/dunesciencealaut0000unse>. (page 9)
- Tasca, C., Rapetti, M., Carta, M. G., & Fadda, B. (2012). Women And Hysteria In The History Of Mental Health. *Clinical Practice & Epidemiology in Mental Health : CP & EMH*, 8, 110–119. <https://doi.org/10.2174/1745017901208010110>. (pages 3, 4)
- Teive, H. A. G., Coutinho, L., Camargo, C. H. F., Munhoz, R. P., & Walusinski, O. (2022). Thomas Willis' legacy on the 400th anniversary of his birth. *Arquivos de Neuro-Psiquiatria*, 80, 759–762. <https://doi.org/10.1055/s-0042-1755278>. (page 4)
- Teive, H. A. G., Germiniani, F., Munhoz, R. P., & Paola, L. d. (2014). 126 hysterical years - the contribution of Charcot. *Arquivos de Neuro-Psiquiatria*, 72, 636–639. <https://doi.org/10.1590/0004-282x20140068.PMID:25098481>. (page 4)
- Tubbs, R. S., Loukas, M., Shoja, M. M., Apaydin, N., Ardalan, M. R., Shokouhi, G., & Oakes, W. J. (2008). Costanzo Varolio (Constantius Varolius 1543–1575) and the Pons Varolli. *Neurosurgery*, 62(3), 734–737. <https://doi.org/10.1227/01.neu.0000317323.63859.2a>. (page 4)
- Varet, V. (2023). Les nouvelles modalités numériques : *blockchain*, Web 3.0, NFT, méta-vers... *Legipresse*, 68(HS1), 59–70. <https://doi.org/10.3917/legip.hs68.0059>. (page 7)
- Varolio, C. (1573). *De nervis opticis nonnullisq : aliis praeter communem opinionem in humano capite obseruatis*. Patavii : apud P. et A. Meiettos fratres. <https://gallica.bnf.fr/ark:/12148/bpt6k325486q>. (page 4)
- Willis, T. (1664). *Cerebri anatome : cui accessit nervorum descriptio et usus*. Londini : Typis Ja. Flesher, impensis Jo. Martyn & Ja. Allestry, apud insigne Campanæ in Coemeterio, D. Pauli. <https://books.google.fr/books/?id=L2xEAAAAcAAJ&pg=PP9#v=onepage&q&f=false>. (page 4)
- Willis, T. (1681). *An Essay of the Pathology of the Brain and Nervous Stock in which Convulsive Diseases are Treated of*. London : Printed by J. B. for T. Dring. <https://quod.lib.umich.edu/e/eebo/A66496.0001.001?rgn=main;view=fulltext>. (page 4)
-

Wright, J. P. (1980). Hysteria and Mechanical Man. *Journal of the History of Ideas*, 41(2), 233–247. <https://doi.org/10.2307/2709458>. (page 3)
