

# Extraction de phrases-clés à partir du corpus Charcot : évaluations quantitative et qualitative

**Ljudmila PETKOVIC<sup>1,2,3,4</sup>**

`prenom.nom@sorbonne-universite.fr`

- <sup>1</sup> Sorbonne Université, Faculté des Lettres, UFR Littératures françaises et comparée, ED III (ED019)
- <sup>2</sup> Sorbonne Université, Centre d'étude de la langue et des littératures françaises (CELLF), UMR 8599
- <sup>3</sup> Sorbonne Université, Observatoire des textes, des idées et des corpus (ObTIC)
- <sup>4</sup> Sorbonne Université, UFR Sociologie et Informatique pour les Sciences Humaines

Séminaire doctoral ObTIC  
SCAI, salle du Conseil  
Paris, le 2 mai 2025



# Table des matières

1. Contributions de Charcot

2. Évaluation quantitative

3. Évaluation qualitative

4. Conclusion

## 1. Contributions de Charcot

## 2. Évaluation quantitative

## 3. Évaluation qualitative

## 4. Conclusion

## Termes inventés par Charcot : référence

nom traditionnel	nom moderne / alternatif
<b>paralysie agitante</b>	maladie de Parkinson
<b>ataxie locomotrice progressive</b>	<i>tabes dorsalis</i>
<b>arthropathies tabétiques</b>	arthropathie de Charcot
<b>trépidation épileptoïde du pied</b>	<i>clonus</i>
<b>sclérose en plaques disséminées</b>	sclérose multiple
<b>sclérose latérale amyotrophique</b>	maladie de Charcot / Lou Gehrig
<b>idée(s) fixe(s), maladie des tics convulsifs</b>	syndrome de Tourette
<b>chorées, athétose</b>	mouvements involontaires
<b>astasia-abasie</b>	incapacité d'être debout / de marcher
<b>atrophie musculaire progressive</b>	maladie Charcot-Marie-Tooth

(WALUSINSKI, 2025; CAMARGO et al., 2023)

≠ termes transmis : **hystérie**  
**épilepsie**  
**hypnose**  
**athétose**

1. Contributions de Charcot

2. Évaluation quantitative

3. Évaluation qualitative

4. Conclusion

# Comparaison des approches d'extraction des phrases-clés

## Critères :

- prise en compte des termes traditionnels ou de leurs synonymes
  - référence : grand dictionnaire terminologique – *La vitrine linguistique*<sup>1</sup>
    - ex. *paralysie agitante* → *maladie de Parkinson*
- recenser le score le plus élevé sur le terme ou sur son synonyme
- méthodes classiques vs. celles de l'état de l'art

NB : termes « imbriqués », contenant eux-mêmes d'autres termes :

- *SEP* : *nystagmus* + *tremblement* + *embarras parole* (bégaiement)

---

1. <https://vitrinelinguistique.oqlf.gouv.qc.ca/>

# Le domaine potentiellement impactant : **hystérie**

Terme	TF-IDF ( <i>TermSuite</i> )	TF-IDF	BM25	PatternRank	Moyenne
<i>maladie de Parkinson</i>	NA	0,2478	0,2397	0,7926	0,3200
<i>ataxie locomotrice progressive</i>	0,1981	0,4114	0,2313	0,7912	0,408
<i>arthropathies tabétiques</i>	0,4424	0,1655	0,337	0,8050	0,4375
<i>trépidation épileptoïde du pied</i>	0,0379	0,2581	0,053	0,7581	0,2768
<i>sclérose en plaques disséminées</i>	NA	0,2935	0,4812	0,7611	0,3840
<i>tremblement</i>	NA	0,2712	0,0213	0,7834	0,2690
<i>nystagmus</i>	NA	0,2142	0,0488	0,7683	0,2578
<i>embarras parole</i>	NA	0,0724	0,9143	0,8159	0,4507
<i>sclérose latérale amyotrophique</i>	NA	0,3287	0,1152	0,7514	0,2988
<i>tics convulsifs</i>	0,0670	0,2273	0,1696	0,8073	0,3178
<i>atrophie musculaire progressive</i>	0,1161	0,2321	0,0797	0,7874	0,3038
<i>aphasie</i>	0,1722	0,345	0,0289	0,7824	0,3321
<i>astisie-abasie</i>	0,1281	0,7022	0,1912	0,7891	0,4527
<i>athétose</i>	NA	0,226	0,0797	0,7910	0,2742
<i>chorées</i>	0,1593	0,1933	0,0213	0,8030	0,2942
<b><i>hystérie</i></b>	<b>0,6892</b>	<b>0,5407</b>	<b>0,0213</b>	<b>0,8194</b>	<b>0,5177</b>
<i>épilepsie</i>	0,0062	0,534	0,0213	0,8170	0,3446
<i>hypnose</i>	0,0311	0,4294	0,0994	0,7955	0,3389
<i>systématisation de l'organisation de la moëlle épinière</i>	NA	0	0	NA	0
<i>localisations cérébrales</i>	NA	0,27	0,0943	0,7493	0,2784

Tab. 1 – Les scores de pertinence pour les termes de référence à partir du corpus «Charcot».

Moyenne pour tous les termes combinés : 0,3280

## Autre domaine potentiellement impactant : syndrome de Tourette

Terme	TF-IDF (TermSuite)	TF-IDF	BM25	PatternRank	Moyenne
<i>maladie de Parkinson</i>	0,05	0,0775	0,333	0,7936	0,3135
<i>ataxie locomotrice progressive</i>	0,32	0,0386	0,4877	0,7431	0,3974
<i>arthropathies tabétiques</i>	0,33	0,0934	0,4928	0,7506	0,4167
<i>trépidation épileptoïde du pied</i>	0,0198	0,1227	0,2919	0,7597	0,2985
<i>sclérose en plaques disséminées</i>	NA	0,178	0,8089	NA	0,2467
<i>tremblement</i>	NA	0,1686	0,0362	0,7683	0,2432
<i>nystagmus</i>	0,0243	0,1326	0,146	0,7474	0,2626
<i>embarras parole</i>	NA	NA	0,0018	NA	0,2341
<i>sclérose latérale amyotrophique</i>	NA	0,044	0,6586	NA	0,1757
<b>tics convulsifs</b>	<b>NA</b>	<b>0,1293</b>	<b>0,8385</b>	<b>0,8331</b>	<b>0,4502<sup>2</sup></b>
<i>atrophie musculaire progressive</i>	0,40	0,1118	0,3489	0,8053	0,4165
<i>aphasie</i>	0,0587	0,2245	0,1334	0,7960	0,3031
<i>astisie-abasie</i>	NA	0,0478	0,3565	0,7375	0,2855
<i>athétose</i>	NA	0,2029	0,274	0,8068	0,3209
<i>chorées</i>	NA	0,1336	0,0701	0,8047	0,2521
<i>hystérie</i>	0,2724	0,3711	0,0442	0,8018	0,3723
<i>épilepsie</i>	NA	0,164	0,0247	0,8199	0,2521
<i>hypnose</i>	0,3543	1	0,2922	0,7738	0,6050
<i>systématisation de l'organisation de la moëlle épinière</i>	NA	NA	NA	0,7550	0,1888
<i>localisations cérébrales</i>	0,43	0,034	0,3017	0,8090	0,3937

Tab. 2 – Les scores de pertinence pour les termes de référence à partir du corpus « Autres ».

Moyenne pour tous les termes combinés : 0,3214

2. Le score le plus élevé parmi les termes **inventés** par Charcot ≠ *hypnose*.



1. Contributions de Charcot

2. Évaluation quantitative

3. Évaluation qualitative

4. Conclusion

# Analyse des concordances des termes médicaux

Analyses effectuées dans TXM.

- on utilise un terme dans le contexte où on cite Charcot
  - terme traditionnel (ex. *paralysie agitante*)
  - synonyme ou terme relevant du champ conceptuel en question
    - ensembles de mots liés à un certain concept (COSTACHESCU, 2024)  
*pied tabétique* → forme particulière d'*arthropathies tabétiques*

## Exemple

```
([word = "paralysie" [word = "agitante" []* [word =  
"Charcot"] | [word = "Charcot" []* [word =  
"paralysie" [word = "agitante"]]) within p
```

- repérer toutes les occurrences dans un paragraphe où *paralysie agitante* et *Charcot* apparaissent dans n'importe quel ordre, séparés par 0 ou plusieurs mots

# Références à Charcot

Terme	Contexte
<i>maladie de Parkinson</i>	[...] <u>paralysie agitante</u> que <b>Charcot</b> a eu raison de dénommer <u>maladie de Parkinson</u>
<i>ataxie locomotrice progressive</i>	[...] arthropathies [...] de l' <u>ataxie locomotrice</u> , [...] signalées, pour la première fois, par M. <b>Charcot</b> . [...] ataxie locomotrice progressive, constatée par douze médecins, parmi lesquels, [...] MM. <b>Charcot</b> [...]
<i>arthropathies tabétiques</i>	[...] mais personne n'avait encore décrit des cas d' <u>arthropathie tabétique</u> , lorsque, en 1808, M. le professeur <b>Charcot</b> publia la première observation d'arthropathie chez un ataxique.
<i>trépidation épileptoïde du pied</i>	[...] trépidation, qui se propage parfois à tous les <u>membres</u> . Ce spasme, [...] peut entraîner à sa suite des rétractions fibro-tendineuses analogues à celles que <b>Charcot</b> a décrites chez l'homme, [...]
<i>sclérose en plaques disséminées</i>	[...] une combinaison de la <u>sclérose en plaques</u> , bien décrite déjà par <b>Charcot</b> et Vulpian [...]

Tab. 3 – Concordance des termes médicaux faisant référence à Charcot – corpus Autres.

# Références à Charcot

Terme	Contexte
<i>tremblement</i>	[...] <b>Charcot</b> présentait, dans son amphithéâtre, pour démontrer les caractères oscillatoires des diverses variétés de <u>tremblements</u> .
<i>nystagmus</i>	SCLÉROSE EN PLAQUES [...] [...], <u>nystagmus</u> , [...]. La difficulté de la résoudre est d'autant plus grande qu'à côté des foyers de sclérose en plaques avec tous les caractères histologiques classiques décrits depuis <b>Charcot</b> , il y a des foyers avec une destruction plus ou moins complète des cylindraxes [...]
<i>embarras parole</i>	Lorsqu'on se trouve en présence d'un malade ayant de l' <u>embarras de la parole</u> [...] A. La réponse à la première proposition n'est nullement embarrassante, si l'on veut se rappeler ces paroles de M. le professeur <b>Charcot</b> : [...]
<i>sclérose latérale amyotrophique</i>	[...] <u>sclérose latérale amyotrophique</u> , maladie découverte par mon illustre maître <b>Charcot</b> .
<i>tics convulsifs</i>	désignée par M. <b>Charcot</b> sous le nom de maladie des <u>tics convulsifs</u>
<i>atrophie musculaire progressive</i>	<b>Charcot</b> et Marie ont décrit la « forme particulière d' <u>atrophie musculaire progressive</u> »
<i>aphasie</i>	Lorsqu'il y a, dit M. le professeur <b>Charcot</b> (1), suppression de la mémoire pour l'articulation des mots, c'est l' <u>aphasie</u> motrice d'articulation ou <u>aphasie</u> de Broca qui se présente.

Tab. 4 – Concordance des termes médicaux faisant référence à Charcot – corpus Autres (suite).

# Références à Charcot

Terme	Contexte
<i>astasia-abasie</i>	[...] forme particulière d'impuissance motrice dont M. P. Blocq a donné la définition suivante « [...], et qu'il a désigné sous le nom expressif d' <b>astasia</b> et d' <b>abasie</b> . C'est là un état morbide sur lequel M. le professeur <b>Charcot</b> est fréquemment revenu dans ses Leçons du mardi [...]
<i>athétose</i>	syndrome désigné par M. W. Hammond sous le nom d' <u>athétose</u> [...] M. <b>Charcot</b> a fait remarquer que cette définition était imparfaite pour les motifs suivants : [...]
<i>chorées</i>	<u>Chorée</u> hystérique ou rythmique. C'est à M. le professeur <b>Charcot</b> que nous devons une exacte description de cet état pathologique.
<i>hystérie</i>	C'est encore à <b>lui</b> [Charcot] que nous devons la connaissance de l' <u>hystérie</u> traumatique [...]
<i>épilepsie</i>	M. <b>Charcot</b> a décrit avec le plus grand soin l' <u>épilepsie</u> partielle d'origine syphilitique [...]
<i>hypnose</i>	Les trois états de l' <u>hypnose</u> décrits par M. <b>Charcot</b> sont devenus classiques, [...]
<i>systématisation de l'organisation de la moëlle épinière</i>	<u>systématisation</u> de la moelle, synthèses [...]. Mais <b>Charcot</b> , on l'a vu, est, par nature, enclin à la synthèse.
<i>localisations cérébrales</i>	Je vous ai montré <b>Charcot</b> , concourant pour la plus grosse part, à l'édification de la doctrine des <u>localisations cérébrales</u> , qui est devenue quelque chose comme la préface d'une psychologie nouvelle.

Tab. 5 – Concordance des termes médicaux faisant référence à Charcot – corpus Autres (fin).

# Analyse des cooccurrences

- quels cooccurrents avec les termes médicaux ciblés?
  - Charcot, Babinski, Necker...
- recensement des résultats pour le cooccurrent : Charcot
- sinon, autre cooccurrent (médecin) avec l'indice le plus élevé

## Exemple

[word = "athétose"]

- liste des cooccurrents pour le terme *athétose*

# Analyse des cooccurrences

Terme	Cooccurent	Fréquence	Co-fréquence	Indice	Distance moyenne
maladie de Parkinson	Charcot	2 968	3	2	3,3
ataxie locomotrice progressive	Charcot	2 968	3	2	5,3
arthropathies tabétiques	Charcot	2 968	4	5	1,5
trépidation épileptoïde du pied	Babinski · Charcot	1 134	8	12	4,8
sclérose en plaques disséminées	Charcot	2 968	7	3	6,3
tremblement	Achard	137	3	2	2,3
nystagmus	Barany	11	3	7	3,3
embarras parole	Chervin	41	8	17	4,5
sclérose latérale amyotrophique	Charcot	2 968	4	3	3,5
tics convulsifs	Charcot	2 968	6	8	5,2
atrophie musculaire progressive	Charcot	2 968	4	3	6,0
aphasie	Charcot	2 968	7	2	4,0
astésie-abasie	Charcot	2 968	2	2	1,5
athétose	Hammond	34	2	4	1,5
chorées	Sydenham	129	63	163	1,1
<b>hystérie</b>	<b>Charcot</b>	<b>2 968</b>	<b>52</b>	<b>19</b>	<b>5,4</b>
épilepsie	Jackson	52	34	78	0,2
hypnose	Braid	567	14	12	4,7
systematisation de l'organisation de la moëlle épinière	NA	NA	NA	NA	NA
localisations cérébrales	Charcot	2 968	9	7	5,3

Tab. 6 – Analyse des cooccurrences des termes médicaux à partir du corpus Autres.

# Analyse des autres cooccurents

Terme	Contexte
<i>trépidation épileptoïde du pied (clonus)</i>	<b>On</b> [Babinski] le désigne alors sous la dénomination de « clonus du pied », « <u>trépidation épileptoïde du pied</u> »
<i>tremblement</i>	D'après <b>Charcot</b> et surtout d'après <b>Achard</b> , ce <u>tremblement</u> aurait de certaines analogies avec le tremblement sénile, [...]
<i>nystagmus</i>	[...] les recherches inspirées par les travaux de <b>Barany</b> sur le <u>nystagmus</u> provoqué indiqueraient une certaine fréquence de troubles labyrinthiques [...]
<i>embarras parole</i>	M. le Dc <b>Chervin</b> , [...] vient de rédiger un nouveau résumé des notions cliniques fondamentales indispensables à connaître sur quelques <u>troubles fonctionnels de la parole</u> et notamment sur le bégaiement.
<i>athétose</i>	[...] nous avons affaire au symptôme désigné par M. W. <b>Hammond</b> sous le nom d' <u>athétose</u> .
<i>chorées</i>	Cependant, la <u>chorée</u> de <b>Sydenham</b> présente quelques particularités sémiologiques que nous allons passer en revue.
<i>épilepsie</i>	Depuis les remarquables travaux de M. Hughlings <b>Jackson</b> sur la forme d' <u>épilepsie</u> à laquelle il a attaché son nom, [...]
<i>hypnose</i>	Selon <b>Braid</b> , l' <u>hypnose</u> est caractérisée par des phénomènes mentaux et physiques, particuliers à cette condition.

Tab. 7 – Concordance des termes médicaux faisant référence à d'autres médecins – corpus Autres.



## Cas ambigu de la *trépidation épileptoïde du pied* (clonus)

**On** le désigne alors sous la dénomination de « clonus du pied », de « trépidation épileptoïde du pied ».

(BABINSKI et BARRÉ, 1934)

[...] l'un d'eux, [...], est connu en France sous le nom de trépidation provoquée, d'épilepsie spinale provoquée. Les auteurs allemands rappellent le phénomène du pied Fussphænomen, ou encore le clonus du pied. Mais c'est là un signe qui appartient à la clinique française. Dès 1863, [...], il était journellement mis à profit dans les services de la Salpêtrière, par M. **Vulpian**, par **moi-même** et par **nos élèves**.

(BRISSAUD et al., 1893)

1. Contributions de Charcot

2. Évaluation quantitative

3. Évaluation qualitative

**4. Conclusion**

# Conclusion

- 1 *PatternRank* : la méthode la plus fiable ( $\neq$  *TermSuite*)
  - capture la sémantique jusqu'aux pentagrammes
- 2 les termes les plus impactants dans les corpus :
  - Charcot : *hystérie*, *astésie-abasie*, *embarras parole*
  - Autres : *hypnose\**, *syndrome de Tourette*, *arthropathies tabétiques*
- 3 absence occasionnelle de cooccurrent Charcot explicable :
  - *épilepsie* : terme créé par J. H. Jackson
  - *hypnose* : terme créé par J. Braid
  - *athétose* : terme créé par W. A. Hammond
  - *chorées* : définition moderne par T. Sydenham
  - *trépidation épileptoïde du pied* : Babinski? Vulpian? Charcot?

Les résultats sont alignés avec les faits historiques.

# Références I



BABINSKI, J. et J.-A. BARRÉ (1934). *Œuvre scientifique*. <https://patrimoine.sorbonne-universite.fr/fonds/item/3486-oeuvres-completes-de-j-m-charcot-tome-4-lecons-sur-les-maladies-du-systeme-nerveux>. Bureaux du progrès médical (Paris) (voir p. 17).



BRISAUD, J.-M. CHARCOT et D. M. BOURNEVILLE (1893). *Œuvres complètes de J. M. Charcot. Tome 4. Leçons sur les maladies du système nerveux*. <https://patrimoine.sorbonne-universite.fr/fonds/item/3486-oeuvres-completes-de-j-m-charcot-tome-4-lecons-sur-les-maladies-du-systeme-nerveux?offset=1>. Bureaux du progrès médical (Paris) (voir p. 17).



CAMARGO, C. H. F., L. COUTINHO, Y. CORREA NETO, E. ENGELHARDT, P. MARANHÃO FILHO, O. WALUSINSKI et H. A. G. TEIVE (2023). Jean-Martin Charcot : the polymath. In : *Arquivos de Neuro-psiquiatria* 81. <https://www.thieme-connect.de/products/ejournals/pdf/10.1055/s-0043-1775984.pdf>, p. 1098-1111 (voir p. 4).



COSTACHESCU, A. (2024). *Les champs linguistiques*. <https://www.lingvistica.ro/imagini/Champs-semantiques.pdf>. Consulté le 19 juin 2025 (voir p. 10).



WALUSINSKI, O. (2025). *Jean-Martin Charcot's Birth Bicentennia*. <http://ishn.org/>. Consulté le 19 juin 2025 (voir p. 4).