

Annexes du document de la séance de regroupement

Contenu

A. Bibliographie	1
B. Information sur notre corpus	3
C. Exemples de requêtes sur notre corpus	4

A. Bibliographie

Adler, S. et Moline, E. (2018). Les noms généraux: présentation. *Langue française*, 2018(2), 5-18.

Aleixandre-Benavent, R., Montalt-Resurecció, V. et Valderrama-Zurián, J. (2014). A descriptive study of inaccuracy in article titles on bibliometrics published in biomedical journals. *Scientometrics*, 101(1), 781-791.

Anthony, L. (2001). Characteristic features of research article titles in computer science. *IEEE Transactions on Professional Communication*, 44(3), 187-194.

Ball, R. (2009). Scholarly communication in transition: The use of question marks in the titles of scientific articles in medicine, life sciences and physics 1966–2005. *Scientometrics*, 79(3), 667-679.

Cheng, S. W., Kuo, C. W. et Kuo, C. H. (2012). Research article titles in applied linguistics. *Journal of Academic Language and Learning*, 6(1), A1-A14.

Cori, M. et David, S. (2008). Les corpus fondent-ils une nouvelle linguistique ? *Langages*, 171, 111-129.

Diers, D. et Downs, F. S. (1994). Colonizing: a measurement of the development of a profession. *Nursing research*, 43(5), 316.

Dillon, J. (1981). The emergence of the colon: an empirical correlate of scholarship. *American Psychologist*, 36, 879-884.

Dillon, J. T. (1982). In Pursuit of the Colon, A Century of Scholarly Progress: 1880–1980. *The Journal of Higher Education*, 53(1).

Flowerdew, J. (2003). Signalling nouns in discourse. *English for specific purposes*, 22(4), 329-346.

Flowerdew, J. (2006). Use of signalling nouns in a learner corpus. *International Journal of Corpus Linguistics*, 11(3), 345-362.

Flowerdew, J. & Forest, R. W. (2015). *Signalling nouns in English*. Cambridge University Press.

Francis, G. (1986). *Anaphoric nouns*. English Language Research, Department of English, University of Birmingham.

Francis, G. (1994). Labelling discourse: an aspect of nominal-group lexical cohesion. In Coulthard, M. ed, (1994), *Advances in written text analysis*, London: Routledge, 83-101.

François, J. et Legallois, D. (2006). Autour des grammaires de constructions et de patterns. *Cahiers du CRISCO*. Université de Caen.

Goodman, R. A., Thacker, S. B. et Siegel, P. Z. (2001). What's in a title? A descriptive study of article titles in peer-reviewed medical journals. *Science*, 24(3), 75-78.

Grant, M. J. (2013). What makes a good title? *Health Information & Libraries Journal*, 30(4), 259-260.

Gustavii, B. (2017). *How to write and illustrate a scientific paper*. Cambridge University Press.

Haggan, M. (2004). Research paper titles in literature, linguistics and science: dimensions of attraction. *Journal of Pragmatics*, 36(2), 293-317.

Halliday, M. A. K. et Hasan, R. (1976). *Cohesion in English*. London: Longman.

Hartley, J. (2005). To attract or to inform: What are titles for? *Journal of technical writing and communication*, 35(2), 203-213.

Hatier, S. (2016). Identification et analyse linguistique du lexique scientifique transdisciplinaire. Approche fouillée sur corpus d'article de recherche en SHS, Thèse de doctorat, Université Grenoble Alpes, 2016.

Hatier, S., Augustyn, M., Tran, T. T. H., Yan, R., Tutin, A. & Jacques, M. P. (2016). French cross-disciplinary scientific lexicon: extraction and linguistic analysis. In *Proceedings of Euralex*, 355-366.

Ho-Dac, L.-M., Jacques, M.-P. & Rebeyrolle, J. (2004). Sur la fonction discursive des titres. Dans S. Porhiel et D. Klingler (éds). *L'unité texte*, Pleyben, Perspectives, 125-152.

Hunston, S. & Francis, G. (1999). *Pattern Grammar. A Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam: Benjamins (Studies in Corpus Linguistics 4).

Huyghe, R. (2018). Généralité sémantique et portage propositionnel: le cas de fait. *Langue française*, 2018(2), 35-50.

Ivanic, R. (1991). Nouns in search of a context: A study of nouns with both open- and closed-system characteristics. *International Review of Applied Linguistics in Language Teaching*, 2, 93-114.

Jacques, T. S. et Sebire, N. J. (2010). The impact of article titles on citation hits: an analysis of general and specialist medical journals. *Journal of the Royal Society of Medicine Short Reports*, 1(1), 1-5.

Jamali, H. R. et Nikzad, M. (2011). Article title type and its relation with the number of downloads and citations. *Scientometrics*, 88(2), 653-661.

Leech, G. N. (2000). Grammars of spoken English: New outcomes of corpus-oriented research. *Language Learning*, 50(4), 675-724.

Legallois, D. (2008). Sur quelques caractéristiques des noms sous-spécifiés. *Scolia*, 23, 109-127.

Mabe, M. A. et Amin, M. (2002). Dr. Jekyll and Dr. Hyde: Author-reader asymmetries in scholarly publishing. *Aslib Proceedings*, 54(3), 149-157.

Merrill, E., & Knipps, A. (2014). What's in a Title?. *The Journal of Wildlife Management*, 78(5), 761-762.

Nagano, R. L. (2015). Research article titles and disciplinary conventions: A corpus study of eight disciplines. *Journal of Academic Writing*, 5(1), 133-144.

Nivard, J. (2010). Les Archives ouvertes de l'EHESS. Récupéré sur *La Lettre de l'École des hautes études en sciences sociales* n°34: <http://lettre.ehess.fr/index.php?5883>

Paiva, C. E., Lima, J. P. da S. N. et Paiva, B. S. R. (2012). Articles with short titles describing the results are cited more often. *Clinics*, 67(5), 509-513.

Rebeyrolle, J., Jacques, M. et Péry-Woodley, M. (2009). Titres et intertitres dans l'organisation du discours. *Journal of French Language Studies*, 19, 269-290.

Roze, C., Charnois, T., Legallois, D., Ferrari, S. et Salles, M. (2014). Identification des noms sous-spécifiés, signaux de l'organisation discursive. Dans *Proceedings of TALN 2014*, 1, 377-388.

Salager-Meyer, F. & Alcaraz Ariza, M. Á. (2013). Titles are "serious stuff": a historical study of academic titles. *Jahr*, 4(7), 257-271.

Sagi, I., & Yechiam, E. (2008). Amusing titles in scientific journals and article citation. *Journal of Information Science*, 34(5), 680-687.

Schmid, H.-J. (2000). *English Abstract Nouns as Conceptual Shells. From Corpus to Cognition*. Berlin: Mouton de Gruyter (Topics in English Linguistics 34).

Schmid, H. J. (2018). Shell nouns in English-a personal roundup. *Caplletra. Revista Internacional de Filologia*, (64), 109-128.

Schwischay, B. (2001). Notes d'exposés sur deux modèles de description syntaxique [Document PDF]. Repéré à <http://www.home.uni-osnabrueck.de/bschwisc/archives/deuxmodeles.pdf>

Soler, V. (2007). Writing titles in science: An exploratory study. *English for Specific Purposes*, 26, 90–102.

Soler, V. (2011). Comparative and contrastive observations on scientific titles written in English and Spanish. *English for Specific Purposes*, 30(2), 124-137.

Subotic, S. & Mukherjee, B. (2014). Short and amusing: The relationship between title characteristics, downloads, and citations in psychology articles. *Journal of Information Science*, 40(1), 115-124.

Swales, J. M. et Feak, C. B. (1994). *Academic Writing for Graduate Students*. Ann Arbor: University of Michigan Press.

Tadros, A. (1994). Predictive categories in expository text. In Coulthard, M. ed, (1994), *Advances in written text analysis*, London: Routledge, 83-96.

Tanguy, L., Rebeyrolle, J. (à paraître). Les titres des publications scientifiques en français : fouille de texte pour le repérage de schémas lexico-syntaxiques.

Townsend, M. A. (1983). Titular Colonicity and Scholarship: New Zealand Research and Scholarly Impact. *New Zealand Journal of Psychology*, 12, 41-43.

Urieli, A. (2013). *Robust French syntax analysis: reconciling statistical methods and linguistic knowledge in the Talismane toolkit*. Toulouse: Doctoral dissertation, Université de Toulouse II-Le Mirail.

Urieli, A. et Tanguy, L. (2013). L'apport du faisceau dans l'analyse syntaxique en dépendances par transitions : études de cas avec l'analyseur Talismane. *Actes de TALN*, Sables D'Olonne.

Winter, E. O. (1977). A clause-relational approach to English texts: a study of some predictive lexical items in written discourse. *Instructional science*, 6(1), 1-92.

Winter, E. O. (1992). The notion of unspecific versus specific as one way of analysing the information of a fund-raising letter. *Discourse description: Diverse linguistic analyses of a fund-raising text*, 131-170.

Yitzhaki, M. (1994). Relation of title length of journal articles to number of authors. *Scientometrics*, 30(1), 321-332.

Yitzhaki, M. (2002). Relation of the title length of a journal article to the length of the article. *Scientometrics*, 54(3), 435-447.

B. Information sur notre corpus

Notre corpus compte 339 687 titres.

Répartition par disciplines

Ordre	Discipline	Titres	% du corpus	% cumulés
1.	0.phys	35538	10.4620 %	10.46 %
2.	1.shs.socio	32228	9.4876 %	19.95 %
3.	1.shs.droit	30517	8.9839 %	28.93 %
4.	1.shs.hist	25764	7.5846 %	36.52 %
5.	NONE	25115	7.3936 %	43.91 %

6.	0.sdv	24638	7.2531 %	51.16 %
7.	1.shs.gestion	23703	6.9779 %	58.14 %
8.	0.info	16932	4.9846 %	63.13 %
9.	1.shs.ling	16028	4.7185 %	67.85 %
10.	1.shs.litt	14938	4.3976 %	72.24 %
11.	1.shs.archeo	13745	4.0464 %	76.29 %
12.	1.shs.edu	9696	2.8544 %	79.14 %
13.	1.shs.scipo	9432	2.7767 %	81.92 %
14.	1.shs.art	8907	2.6221 %	84.54 %
15.	1.shs.phil	8450	2.4876 %	87.03 %
16.	0.sde	7621	2.2435 %	89.27 %
17.	1.shs.infocom	7333	2.1588 %	91.43 %
18.	1.shs.anthro	7095	2.0887 %	93.52 %
19.	1.shs.archi	4761	1.4016 %	94.92 %
20.	0.sdu	3842	1.1310 %	96.05 %
21.	0.scco	3141	0.9247 %	96.98 %
22.	0.math	3092	0.9102 %	97.89 %
23.	0.chim	2836	0.8349 %	98.72 %
24.	1.shs.psy	2654	0.7813 %	99.51 %
25.	1.shs.geo	1149	0.3383 %	99.84 %
26.	0.qfin	433	0.1275 %	99.97 %
27.	1.shs.autre	99	0.0291 %	100.00 %

On note qu'au niveau des domaines, le corpus n'est pas équilibré. Il faut donc toujours procéder par fréquences relatives pour comparer une caractéristique d'une discipline à l'autre.

C. Exemples de requêtes sur notre corpus

Notre code en Python sert essentiellement à faire des statistiques sur notre corpus et le filtrer.

```
stat('domain')
```

Produit un comptage des titres selon la discipline des titres. Le résultat est un dictionnaire où la clé est la discipline et la valeur le nombre de titre dans cette discipline.

```
stat(('nb_parts', 'nb_segments'))
```

Produit un comptage des titres selon les combinaisons des valeurs possibles pour le nombre de parties et le nombre de segments. Le résultat est un dictionnaire où la clé est un tuple constitué d'une combinaison existante de valeurs des deux dimensions, par exemple 1 partie, 2 segments, et la valeur le nombre de titre correspondant à cette combinaison, le nombre de titres ayant 1 partie et 2 segments.

```
count({'nb_parts' : 1, 'nb_segments' : 2})
```

Compte le nombre de titre ayant une partie et deux segments.

```
t12 = select({'nb_parts' : 1, 'nb_segments' : 2})
```

Création d'un sous-corpus composé des titres ayant une partie et deux segments. On peut ensuite utiliser les requêtes stat et count sur celui-ci via une variable globale qui contient le corpus courant.