

*Étude comparative des relations franco-
allemandes dans les trois journaux
Le Monde, Le Figaro et Süddeutsche Zeitung.*

Une approche textométrique

Jane Wottawa LIUM, Le Mans Université

Nantes, le 13 et 14 juin 2024



Introduction

- Projet « Amorçage transversal » MSH Ange-Guépin:
Penser les relations franco-allemandes, collaboration entre Henning Fauser (CRINI, Nantes Université) et Jane Wottawa (LIUM, Le Mans Université)
- Encadrement de deux étudiants de Lettres modernes en *Projet tutoré* (20h chacun, L3): Théo Lamballe et Louisa Provost
 - Introduction à la *textométrie* et au logiciel TXM
 - Création de corpus homogènes pour les journaux *Süddeutsche Zeitung*, *Le Monde*, *Le Figaro*
 - Mener quelques analyses afin d'exploiter le sous-corpus

Qu'est-ce que la *textométrie*?

- Aussi appelée *lexicométrie* ou *statistique lexicale* ou *textuelle*
 - Permet d'évaluer la richesse du vocabulaire employé dans un texte ou un corpus (liste des fréquences)
 - Regarder les mots dans leur contexte immédiat (concordances)
 - Rendre compte des co-occurrences à travers tous les documents qui constituent le corpus (co-occurrences)
- Un outil qui permet de trouver les **systématicités** dans un corpus
- Dépend de la langue – analyses séparées pour l'allemand et le français

Questions de recherche

- I. Dans une même période, observons nous des différences entre les termes employés dans les trois journaux (*Süddeutsche Zeitung*, *Le Monde*, *Le Figaro*) ?

- II. Quelles informations la *textométrie* nous donne-t-elle quant aux imaginaires mobilisées en Allemagne et en France ?

Corpus

	<i>Süddeutsche Zeitung</i> (Jane Wottawa)	<i>Le Monde</i> (Louisa Provost)	<i>Le Figaro</i> (Théo Lamballe)	Total
Période	2012-2022 avec une surreprésentation de 2012 (Merkel, Hollande, conseil européen en décembre 2012)			-
Nombre de textes	20	22	21	63
Nombre de mots	19 842	21 527	21 376	62 745
Orientation	Centre-gauche	Centre-gauche	Droite / centre-droit	-

- 3 journaux : 1 allemand, 2 français ; 2 centre-gauche, 1 centre-droit
- Méthode : exploitation avec le logiciel TXM
- Hypothèse : selon le pays et selon l'orientation politique du journal, les textes présentent des sensibilités différentes.

Fréquence de mots

- La fréquence de mots nous informe sur les termes importants dans chacun des corpus
- Méthode : les lemmes lexicaux les plus fréquents des trois journaux ont été relevés
- Une liste avec les mots le plus fréquents a été créé à partir des trois corpus: *Europe, Allemagne, allemand, France, Merkel, européen, ensemble, UE, président, français, Berlin, euro, Paris, Hollande, franco-allemand*
- En somme: 15 mots dont 5 adjectifs (les autres mots sont des noms)

Lemma ; lemme	<i>Süddeutsche Zeitung</i>	<i>Le Monde</i>	<i>Le Figaro</i>
Europa ; Europe	120	63	70
Deutschland ; Allemagne	69	38	38
deutsch ; allemand	61	49	33
Frankreich ; France	59	52	81
Merkel	58	46	24
europäisch ; européen	57	143	69
gemeinsam ; ensemble	39	9	4
EU ; UE	37	40	15
Präsident ; président	33	41	39
französisch ; français	31	53	74
Berlin	22	56	25
Euro	15	56	36
Paris	18	49	31
Hollande	<i>mot</i> 14	23	49
deutsch-französisch ; franco-allemand	16	7	37

Conclusion fréquence de mots

- Les journaux centre-gauche ont plus d'occurrences mentionnant l'**Europe**, le journal centre-droite a comme lemme le plus fréquent **France**
- En politique, les journaux français font plus souvent référence à la **ville** où siège le gouvernement que le journal allemand (*Paris, Berlin*)
- Le sous-corpus du *Figaro* présente plus fréquemment le lemme **franco-allemand** que les deux autres journaux
- Chaque journal parle fréquemment de **son propre pays** (Allemagne, France)

Analyse des co-occurrences

- Co-occurrence: *la co-occurrence est la **co-présence** ou **présence simultanée** de **deux unités linguistiques** (deux mots par exemple ou deux codes grammaticaux) au sein d'un même contexte linguistique (le paragraphe ou la phrase par exemple, ou encore une fenêtre arbitraire) (Mayaffre, 2007, p. 92)*
- Méthode: utilisation des 15 mots les plus fréquents afin de déterminer leurs co-occurrences dans les trois sous-corpus
- Hypothèse : les co-occurrences nous informent davantage sur le sens du mot pivot et ses nuances dans les trois sous-corpus

Co-occurrences

Lemma ; lemme	<i>Süddeutsche Zeitung</i>	<i>Le Monde</i>	<i>Le Figaro</i>
Europa ; Europe	120	63	70

Occ	Fréq	CoFréq	Indice	Occ	Fréq	CoFréq	Indice	Occ	Fréq	CoFréq	Indice	DistMoy
Zukunft	7	4	2.4171	menace	6	3	2.4	vers	12	4	2.2803	1.75
bedeuten	4	3	2.3166	champion	3	2	2.0	sud	3	3	3.5968	2.0
neu	44	15	4.4092	continent	7	3	2.2	besoin	8	4	3.0409	2.0
für	117	24	2.7527	élargissement	3	2	2.0	menace	2	2	2.3976	6.0
brauchen	11	5	2.3858	faible	2	2	2.4	mobiliser	2	2	2.3976	6.0
sondern	25	9	3.0512									
Idee	13	8	4.7991									
Herausforderung	5	4	3.1832									
Frieden	10	5	2.6082									
müssen	64	19	4.4522									
Art	8	5	3.1793									
Solidarität	7	4	2.4171									

- polarisation entre la SZ et les journaux français

Co-occurrences

Lemma ; lemme	Süddeutsche Zeitung	Le Monde	Le Figaro
Deutschland ; Allemagne	69	38	38

Occ	Fréq	CoFréq	Indice	Dist	Occ	Fréq	CoFréq	Indice	Dist	Occ	Fréq	CoFréq	Indice	DistMoy
Wirtschaftskongress	2	2	2.3445	1.5	basculer	2	2	2.934	3.5	France	81	12	4.7033	3.9166667
Frankreich	59	26	14.9175	1.76	France	52	15	9.9208	4.0	entre	38	6	2.7608	5.1666665
Export	6	3	2.2833	6.0	côté	23	4	2.1479	5.0	nazi	3	2	2.4619	3.0
sozialistisch	2	2	2.3445	7.5	Nord	3	2	2.4668	5.5	armée	4	2	2.1709	2.5
Bundesrepublik	2	2	2.3445	8.5	sommet	7	3	2.9025	6.6666665					

économie, politique

géo-politique

conflit

- co-occurrence commune: *France*

Co-occurrences

Lemma ; lemme	<i>Süddeutsche Zeitung</i>			<i>Le Monde</i>			<i>Le Figaro</i>				
deutsch ; allemand	61			49			33				
Occ	Fréq	CoFréq	Occ	Fréq	CoFréq	In	Occ	Fréq	CoFréq	Indice	DistMoy
Verbrechen	2	2	presse	7	3		chancelière	15	5	4.1927	0.4
Seite	8	3	chancelière	17	9		gouvernement	17	5	3.901	1.2
Französisch	31	9	finances	14	4		droit droite	6	3	3.2624	1.0
Unternehmen	13	4	Olaf	3	2		français	74	8	2.7021	4.625
			Angela	21	5		Moscovici	5	2	2.0484	3.5
			Scholz	2	2	2.7174		3.0			
			ministre	46	7	2.439		3.5714285			
			Wolfgang	4	4	5.4365		4.25			

- SZ : plusieurs thèmes sont abordés dont l'économie
- Les journaux français présentent plutôt des aspects politiques
- *Le Monde* : mention de la presse et des finances

Co-occurrences

Lemma ; lemme	<i>Süddeutsche Zeitung</i>	<i>Le Monde</i>	<i>Le Figaro</i>
Frankreich ; France	59	52	81

Occ	Fréq	CoFréq	Indice	Occ	Fréq	CoFréq	Int	Occ	Fréq	CoFréq	Indice	DistMoy
Großbritannien	13	4	2.26	Hollande	23	5		doter	5	4	3.8584	2.25
Veto	2	2	2.465	Allemagne	38	15		plan	19	6	2.7285	2.8333333
Deutschland	69	26	14.379	Bruxelles	36	6		rang	4	3	2.8205	3.3333333
Solidarität	7	3	2.232	Espagne	5	3		Allemagne	38	12	4.9585	3.9166667
Amtszeit	3	2	2.00	François	16	4		frappe	2	2	2.265	4.0
Diskussion	3	2	2.006		8.0			Grande-Bretagne	2	2	2.265	6.5

expliquer la France

relations

recevoir et imposer

Lemma ; lemme	<i>Süddeutsche Zeitung</i>	<i>Le Monde</i>	<i>Le Figaro</i>
Europa ; Europe	Idéologie, politique	Danger, géographie	Danger, géopolitique
Deutschland ; Allemagne	Politique , économie	Politique , géographie	Danger, politique
deutsch ; allemand	Danger, économie	Politique, économie	Politique
Frankreich ; France	Relations	Relations	Argent, relations
Merkel	Politique, relations	Relations déséquilibrées	Relations
europäisch ; européen	Communauté, politique	Argent, politique	Argent, politique
gemeinsam ; ensemble	Défense, économie	<i>Non représentatif</i>	<i>Non représentatif</i>
EU ; UE	Relations, argent	Politique	Opportunité
Präsident ; président	France	Relations	France
französisch ; français	Politique, économie	Argent	Idéologie
Berlin	Politique, défense	Argent, économie	Relations, argent
Euro	Argent	Argent , relations	Argent
Paris	Relations	Relations , représentation	Relations
Hollande	Politique	Relations	Relations, politique
deutsch-französisch ; franco-allemand	Relations	<i>Non représentatif</i>	Relations

Conclusion co-occurrences

- Il y a **consensus** de champs lexicaux pour les 6 termes pivots suivants:
Allemagne, France, Merkel, européen, Euro, Paris
- La perception de l'**Europe** diffère grandement entre le journal allemand et les deux journaux français
- La *Süddeutsche Zeitung* met plus en valeur l'**économie** que les deux journaux français, *le Figaro* ne mentionne pas cet aspect
- Les journaux français parlent plus des **aspects monétaires** notamment en lien avec l'Europe

Conclusions générales

- I. Dans une même période, observons nous des différences entre les termes employés dans les trois journaux?
 - Oui, il existe des différences mais aussi des lieux communs. Certains termes ne sont pas du tout associés au même référentiel (*i.e.* Europe) → fait partie de la *Perzeption* ?

- I. Quelles informations la *textométrie* nous donne-t-elle quant aux imaginaires mobilisées côté allemand et côté français?
 - La fréquence de mots nous indique un ordre d'importance de certains termes employés.
 - Les co-occurrences peuvent guider des analyses qualitatives ensuite.

References

Citées

- Gledhill, C., Patin, S., & Zimina, M. (2017). Lexico-grammaire et textométrie: identification et visualisation de schémas lexico-grammaticaux caractéristiques dans deux corpus juridiques comparables en français. *Corpus*, (17).
- Heiden S., Magué J., & Pincemin B. (2010). TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement. In *JADT 2010 : 10th International Conference on the Statistical Analysis of Textual Data* (pp. 12 p.). Rome, Italie.
- Mayaffre, D. (2007). L'entrelacement lexical des textes. Cooccurrences et lexicométrie. In *Journées de Linguistique de Corpus* (No. 3, pp. 91-102).

Consultées

- Buhler, T., Bendinelli, M., Lethier, V., & He, Y. (2018). La textométrie pour révéler évolutions et oppositions dans les discours des politiques urbaines. Le cas des PDU français (2000-2015). *Flux*, (3), 93-110.
- Leblanc, J. M. (2015). Proposition de protocole pour l'analyse des données textuelles: pour une démarche expérimentale en lexicométrie. *Nouvelles perspectives en sciences sociales*, 11(1), 25-63.
- Mayaffre, D., Pincemin, B., & Poudat, C. (2019). Explorer, mesurer, contextualiser. Quelques apports de la textométrie à l'analyse de discours. *Langue française*, (203), 101-115.
- Pengam, M., Jackiewicz, A., & Marchand, P. (2022, July). Saisir par la textométrie la notion de radicalisation dans des corpus institutionnels hétérogènes. In *16th International Conference on Statistical Analysis of Textual Data (JADT 2022)*.