

Colloque international JADT 2022 : 16th International Conference on Statistical Analysis of Textual Data

6-8 juillet 2022

Communication acceptée

Saisir par la textométrie la notion de radicalisation dans des corpus institutionnels hétérogènes

Manon Pengam¹, Agata Jackiewicz², Pascal Marchand³

¹Université Paul-Valéry Montpellier 3 – manon.pengam@univ-montp3.fr

²Université Paul-Valéry Montpellier 3 – agata.jackiewicz@univ-montp3.fr

³Université Toulouse 3 Paul Sabatier – pascal.marchand@iut-tlse3.fr

Résumé

Dès le début des années 2010, la problématique de la radicalisation djihadiste a généré en France une importante production de discours institutionnels et une série de plans d'actions politiques destinés à éviter la commission d'actes terroristes sur le territoire national.

Notre travail, situé au confluent de l'analyse de discours et de la statistique textuelle, propose de comparer deux types de discours institutionnels sur la radicalisation : des déclarations publiques de l'exécutif (2013-2018), issus de Pengam (2021) et des rapports provenant des commissions d'enquêtes parlementaires (2013-2017), issus de Marchand (2017), respectivement 2 et 1 millions d'occurrences.

En première lecture, ces deux corpus possèdent des orientations argumentatives différentes et des caractéristiques énonciatives propres. Les rapports d'enquêtes parlementaires prescrivent des recommandations qui émanent du législateur mais aussi d'acteurs de la société civile (syndicalistes, témoins, avocats...) auditionnés en raison de leur expertise à l'égard de la problématique considérée. Les déclarations publiques de l'exécutif (ministres, Président de la République) imposent, quant à elles, une vision homogène et consensuelle des problèmes publics, nivelant la diversité des points de vue exprimés.

Dans un contexte d'urgence sociale et politique liée aux attentats terroristes, dans quelle mesure – et si oui comment – les recommandations émises par les commissions d'enquête sont-elles reprises par l'action publique ? Plus généralement, comment se construit et évolue la représentation publique de la radicalisation djihadiste ?

Pour répondre à ces questions, nous nous proposons de réaliser une étude contrastive de ces deux ensembles discursifs, (i) par le recours à des outils d'exploration et de calcul statistiques, (ii) par l'utilisation de modèles linguistiques.

1. (i) Sur le plan méthodologique, nous cherchons à caractériser le degré de proximité thématique de ces corpus au moyen des relations entre classifications lexicales (Reinert 1983, 1990) à l'aide des logiciels IRaMuTeQ et R, selon trois modes de calcul : une analyse factorielle des correspondances sur le tableau lexical complet (Lebart et Salem, 1994), une mesure des distances textuelles avec l'indice de Labbé (Labbé et Labbé, 2003), et une utilisation des types généralisés (Lamalle et Salem, 2002 ; Marty, Marchand et Ratinaud, 2013).
2. (ii) Sur le plan linguistique, il s'agit de saisir, par l'étude de la matérialité discursive (cooccurrences, patrons lexico-syntaxiques, métaphores prototypiques comme *X est le terreau de la radicalisation...*), le processus d'institutionnalisation d'une notion complexe dont les manifestations (référentielles,

énonciatives, argumentatives) sont susceptibles d'évoluer et de diverger, notamment selon les effets pragmatiques visés. Nous utilisons pour cela IRaMuTeQ et TXM.

Les premières explorations par classification hiérarchique descendante sur segments de textes montrent d'ores-et-déjà une similarité temporelle entre les thématiques abordées par les deux corpus. Par exemple, en 2015, Internet apparaît avec une force particulière comme une des causes de l'engagement djihadiste, tandis qu'en 2017, c'est l'environnement carcéral qui sera perçu comme un vecteur important de la radicalisation. Notre étude permettra de dresser une «frise» temporelle contrastive, mettant précisément en évidence l'évolution longitudinale et la circulation des causes explicatives de la radicalisation dans les deux sources de discours.

Les analyses à venir seront complétées de focus sur des énoncés et marqueurs linguistiques typiques de la notion de radicalisation, telle qu'analysée dans Pengam (2021) : marqueurs d'élaboration et de cadrage référentiel, marqueurs processuels de progression et de franchissement d'étapes, marqueurs causaux...

Plus largement, nous souhaitons par ce travail qui s'inscrit dans une démarche générale d'analyse des notions socio-politiques complexes et sensibles (Jackiewicz et al., 2019 ; Jackiewicz et Pengam, 2020 ; Ratinaud et Marchand, 2012) : (i) proposer une méthode d'exploration transversale mêlant approche textométrique et approche linguistique d'inspiration sémantico-discursive ; (ii) fournir une contribution à l'analyse des problèmes publics au sens de Neveu (2017), tels qu'ils sont représentés par la parole institutionnelle (Née, Oger et Sitri, 2017 ; Krieg-Planque, 2012 ; Oger et Ollivier-Yaniv, 2003).

Références

- Jackiewicz, A. et Pengam, M. (2020). Un modèle pour l'étude des nominations émergentes. Notion de repérage pour saisir les modalités d'ajustement sémantique et discursif. *SHS Web of Conferences*, 78, *Congrès Mondial de Linguistique Française 2020*. <https://doi.org/10.1051/shsconf/20207812004>
- Jackiewicz, A., Bebeskina, N., Cassier, M., Frontini, F., Halftermeyer, A., Longhi, J., Luxardo, G., Nouvel D. (2019). *Vers une ontologie de la nomination et de la référence dédiée à l'annotation des textes. ToTH 2019*, Chambéry.
- Krieg-Planque, A. (2012). *Analyser les discours institutionnels*. Armand Colin. Labbé, C. et Labbé, D. (2003). La distance intertextuelle. *Corpus*, (2) : 95-117.
- Lamalle, C. et Salem, A. (2002). Types généralisés et topographie textuelle dans l'analyse quantitative des corpus textuels. *Sixièmes journées d'analyse des données textuelles*, Saint-Malo, 403-412.
- Lebart, L. et Salem, A. (1994). *Statistiques textuelles*. Dunod.
- Marchand, P. (2017). « La fabrique parlementaire du discours sur la “radicalisation” : politiques, acteurs, experts », dans Baygert Nicolas, Durin Esther, Maas Élise et Nicolas Loïc (dirs.), *Communiquer (sur) la radicalité*, Les Cahiers PROTAGORAS, 4, 2017, 30-46.
- Marty, E., Marchand, P. et Ratinaud, P. (2013). Les médias et l'opinion : éléments théoriques et méthodologiques pour une analyse du débat sur l'identité nationale. *Bulletin de Méthodologie Sociologique*, 117(1), 46-60.
- Née, É., Oger, C. et Sitri, F. (2017). Le rapport : opérativité d'un genre hétérogène. *Mots. Les langages du politique*, 114, 9-24.
- Neveu, É. (2017). L'analyse des problèmes publics. *Idées économiques et sociales*, 190(4), 6-19.
- Oger, C. et Ollivier-Yaniv, C. (2003). Analyse du discours institutionnel et sociologie compréhensive : vers une anthropologie des discours institutionnels. *Mots. Les langages du politique*, 71, 125-145. <https://doi.org/10.4000/mots.8423>
- Pengam, M. (2021). *Pour une modélisation linguistique de la radicalisation. Étude de discours institutionnels et de discours du travail social*. [thèse de doctorat, Université Paul-V aléry Montpellier 3].
- Ratinaud, P. et Marchand, P. (2012). Recherche improbable d'une homogène diversité : le débat sur l'identité nationale. *Langages*, 187, 93-107.

Reinert, M. (1983). Une méthode de classification descendante hiérarchique : application à l'analyse lexicale par contexte. *Les Cahiers de l'analyse des données*, 3, 187-198.

Reinert, M. (1990). ALCESTE: Une méthodologie d'analyse des données textuelles et une application. *Bulletin de Méthodologie Sociologique*, 26, 24-54.