



HAL
open science

Analyse et cartographie des sentiments dans des récits de vie de migrants

Catherine Dominguès, Serge Weber, Carmen Brando, Laurence Jolivet,
Marie-Dominique van Damme

► **To cite this version:**

Catherine Dominguès, Serge Weber, Carmen Brando, Laurence Jolivet, Marie-Dominique van Damme. Analyse et cartographie des sentiments dans des récits de vie de migrants. Spatial Analysis and GEomatics 2017, INSA de Rouen, Nov 2017, Rouen, France. hal-01649150

HAL Id: hal-01649150

<https://hal.science/hal-01649150v1>

Submitted on 27 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Analyse et cartographie des sentiments dans des récits de vie de migrants

Catherine Dominguès¹, Serge Weber², Carmen Brando³,
Laurence Jolivet¹, Marie-Dominique Van Damme¹

1. Université Paris-Est, IGN/SR, COGIT,
73 avenue de Paris, F-94165 Saint-Mandé cedex, France
prenom.nom@ign.fr
2. Université Paris-Est Marne-la-Vallée, laboratoire Analyse comparée des pouvoirs
5 boulevard Descartes, F-77454 Marne-la-Vallée Cedex 2, France
prenom.nom@u-pem.fr
2. Ecole des hautes études en sciences sociales (EHESS)
54 boulevard Raspail, F-75006 Paris, France
prenom.nom@ehess.fr

RESUME. Cet article présente le projet MATRICIEL (Projet Exploratoire Premier soutien) subventionné par le CNRS et l'Université Paris-Est. Le projet est fondé sur l'analyse et la cartographie d'un corpus de récits de vie de républicains espagnols ; il s'intéresse à la migration sous l'angle des lieux décrits par les migrants dans les récits de leur vie. Le corpus a été transcrit puis annoté en lieux (noms propres et lieux génériques) et sentiments (polarité) afin de constituer des corpus d'apprentissage et de validation pour la reconnaissance des lieux et des sentiments associés. Des modèles d'apprentissage automatique pour la reconnaissance automatique des noms de lieux génériques ont été entraînés. Les sentiments associés aux lieux seront inférés à partir d'un lexique de sentiments analysé en contexte. La représentation vise une cartographie des informations sensibles liées aux lieux.

ABSTRACT. This article introduces the project MATRICIEL (exploratory project, first support) funded by the CNRS (France's national scientific research center) and the University Paris-Est. The project is based on the analysis and mapping of Spanish Republicans' life-stories; it addresses migration issue from the places described by migrants in their narratives. The corpus was transcribed, then annotated with places (proper nouns and generic place nouns) and sentiments (polarity) in order to form learning corpus and test corpus to automatically retrieving places and associated sentiments. Machine learning models to generic place noun identification were trained. Sentiments associated to places will be inferred from sentiment lexicon analyzed in context. The representation aims to map sensitive data linked to places.

MOTS-CLES : recherche d'entité nommée spatiale, analyse de sentiment, nom de lieu générique, corpus oral, récit de vie, carte sensible, migration, républicain espagnol, RAHMI

KEYWORDS: Named Entity Recognition (NER), sentiment analysis, generic location, spoken corpus, life story, sensitive map, migration, Spanish Republican, RAHMI

1. Introduction

La migration est un phénomène ancien mais constitue un sujet d'intérêt toujours renouvelé comme en témoignent l'actualité immédiate et son traitement médiatique, l'apparition de nouveaux types de migrations (par exemple les migrations climatiques) ainsi que les productions universitaires sur le sujet. Dans ce contexte, le projet MATRICIEL s'intéresse à la migration sous l'angle des lieux décrits par les migrants dans les récits de leur vie. Rendre compte des sentiments attachés à ces lieux contribue à explorer le statut de la trace dans la mémoire de l'immigration comme patrimoine collectif (Musée de l'histoire de l'immigration, 2014).

2. État de l'art

La question du sensible en sciences sociales a fait l'objet de nombreuses interrogations autour du problème, dans le travail empirique, de la saisie et de la restitution des émotions (des témoins, du chercheur) obligeant à un travail de réflexivité sur les conditions de l'interaction de recherche (Cefai, 2010). Les « géographies émotionnelles » se sont multipliées notamment dans les terrains dits « sensibles », où les émotions compliquent le processus d'objectivation (Pian, 2012). Cartographier les voyages des migrants et réfugiés a donné lieu à des expérimentations représentationnelles laissant une place au temps perdu, à l'immobilisation contrainte (Darley *et al.*, 2013). Des outils de recherche automatique d'information : lieux, dates, sentiments (Edwardes et Purves, 2007 ; Ballatore et Adams, 2015), de représentation cartographique (Caquard et Fiset, 2014) ont été construits. La cartographie telle qu'elle est habituellement pratiquée s'avère insuffisante pour rendre compte du ressenti des rapports sociaux, de même que la représentation strictement visuelle d'informations (Lorimer, 2005), en particulier pour la migration. Ce projet vise à explorer la narration de l'expérience vécue de l'immigration tenant compte de la dimension émotionnelle et langagière, afin de dégager des pistes méthodologiques de représentation à la croisée entre le *mapping* et le *sounding* (Canova *et al.*, 2014), que certains ont réussi à restituer par l'écriture et la réalisation de fictions ethnographiques (Mai, 2014). Nous proposons dans ce projet de croiser des méthodes liées au TALN, à la géographie et à la cartographie pour rendre compte de l'importance des sentiments associées aux lieux dans l'expérience migratoire et pour placer la parole des migrants au premier plan de l'analyse.

3. Méthode proposée

Le projet MATRICIEL a pour corpus d'étude dix-sept récits oraux de résistants espagnols, recueillis en 2009 par le Réseau aquitain pour l'histoire et la mémoire de l'immigration (RAHMI) afin de documenter le parcours des résistants espagnols ayant combattu puis fui le franquisme en Espagne, pour s'installer finalement en France. Ces récits (durée 18 h 30 mn) ont été transcrits puis annotés manuellement

en lieux et sentiments. Le lieu est ici compris dans un sens large : le lieu désigné par un toponyme répertorié dans un dictionnaire de toponymes (le plus souvent un nom propre, par exemple *France*) mais aussi celui désigné par un nom générique, éventuellement précisé par un nom propre, pour évoquer les lieux d'arrivée, de transit, d'asile, etc. comme *camp d'internement*, *camp de Gurs* ou des objets localisés comme *convoi*, *dépôt d'armes* (Brando *et al.* 2016). Le sentiment est souvent défini en TALN comme une polarité (positive ou négative) qui est attachée à un segment de texte considéré en contexte, comme : *ami*, *fascisme*, *engagé*, *accueillir*, puis ensuite propagée sur un/des lieu(x) comme dans : *le fascisme s'installe en Europe Italie et en Espagne* où *fascisme* est annoté négativement, et des liens sont posés entre *fascisme* et chacune des entités *Europe*, *Italie* et *Espagne*.

L'exploration de ce corpus s'appuie sur des méthodes et des outils de traitement automatique des langues, intégrés dans l'outil GATE¹. L'identification et la localisation des lieux noms propres sont effectuées à l'aide de gazetiers construits à partir des bases de données BD TOPO² et Geonames³, et le module ANNIE⁴ de GATE. L'identification des lieux génériques et des objets localisés est fondée sur un modèle entraîné à partir d'un extrait annoté du corpus et l'outil d'apprentissage automatique Stanford-NER⁵. L'annotation automatique en sentiments (polarité) est en cours de réalisation grâce à l'adaptation au français des travaux et outils de (Andreevskaia et Bergler, 2007) et (Ozdemir et Bergler, 2015). Les modèles produits sont évalués en termes de rappel, précision, F-mesure par rapport au corpus de référence.

Les représentations cartographiques visent à représenter les lieux évoqués dans les récits de vie ainsi que des informations associées, dont les sentiments inférés. L'objectif de ces représentations est de faciliter l'exploration des récits et de les valoriser sous la forme de cartes du sensible qui intègrent des polarités (Griffin et McQuoid, 2012). Dans ce contexte, les enjeux concernent la représentation de la temporalité (chronologie, durée, simultanéité) des sentiments associés aux lieux et celle des perceptions multiples (exprimées dans un récit ou plusieurs récits) d'un lieu ; l'incertitude sur la localisation d'un lieu et sur son emprise (*le camp de Gurs*) ; et enfin, la prise en compte de la métonymie (par exemple, un nom de lieu *Espagne* désignant un sous-ensemble comme la ville ou le village, lieux de naissance ou de séjour du locuteur).

Remerciements

Ce projet, subventionné par projet subventionné par le programme PEPS CNRS-UPE, n'aurait pu se mettre en place sans la confiance du RAHMI (Réseau aquitain pour l'histoire et la mémoire de l'immigration).

¹ <https://gate.ac.uk/>

² <http://professionnels.ign.fr/bdtopo>

³ <http://www.geonames.org/>

⁴ <https://gate.ac.uk/ie/annie.html>

⁵ <https://nlp.stanford.edu/software/CRF-NER.shtml>

Références

- Andreevskaia A., Bergler S. (2007). [CLaC and CLaC-NB: Knowledge-based and corpus-based approaches to sentiment tagging](#). In: *Proceedings of SemEval-2007: 4th International Workshop on Semantic Evaluations at ACL 2007*, Prague <http://www.aclweb.org/anthology/S/S07/S07-1022.pdf>
- Ballatore A., Adams B. (2015). Extracting Place Emotions from Travel Blogs, AGILE 2015, Lecture Notes in Geoinformation and Cartography, Lisbon, June 2015, pp. 1–5.
- [Brando, C., Dominguès C., Capeyron M. \(2016\). Evaluation of NER Systems for the Recognition of Place Mentions in French Thematic Corpora](#), *Proceedings of the 10th Workshop on Geographic Information Retrieval*, pp. 7:1-7:10, GIR '16, ACM, [doi:10.1145/3003464.3003471](https://doi.org/10.1145/3003464.3003471)
- Canova N., Bourdeau P., Soubeyran O. (dir.). (2014). *La petite musique des territoires : art, espace et sociétés*, Paris, CNRS Éditions.
- Caquard S., Fiset J.P. (2014). How Can We Map Stories? A Cybercartographic Application for Narrative Cartography, *The Journal of Maps* 10(1):18-25 [doi:10.1080/17445647.2013.847387](https://doi.org/10.1080/17445647.2013.847387)
- Cefai D. (dir.). (2010). *L'engagement ethnographique*, Paris, Ehess.
- Darley M., Lancelevée C., Michalon B. (2013). *Où sont les murs ? Penser l'enfermement en sciences sociales*, Paris, L'Harmattan, p. 148.
- Edwardes A.J., Purves, R.S. (2007). A theoretical grounding for semantic descriptions of place. *Proceedings of 7th Intl. Symposium on Web and Wireless Geographical Information Systems (W2GIS)*. Cardiff, UK, 28-29th Nov. 2007,(4857). Ware, M. and Taylor, G.(eds), p. 106-121.
- Griffin A.L., McQuoid J. (2012). At the Intersection of Maps and Emotion: The Challenge of Spatially Representing Experience. *Kartographische Nachrichten*, 62(6). 291–299.
- Lorimer H. (2005). Cultural geography : the busyness of being 'more-than-representational', *Progress in Human Geography*, vol. 29, p. 83-94.
- Mai N. (2014). Between Embodied Cosmopolitanism and Sexual Humanitarianism : the Fractal Mobilities and Subjectivities of Migrants Working in the Sex Industry, in Baby-Collins, V. and Anteby, L. (eds) *Borders, Mobilities and Migrations, Perspectives from the Mediterranean in the 21st Century*, Brussels : Peter Lang, p. 175-192.
- Musée de l'histoire de l'immigration (2014). *Guide de la Galerie des dons*, Paris, Etablissement public du palais de la Porte Dorée, Musée national de l'histoire de l'immigration
- Ozdemir C., Bergler S. (2015). A Comparative Study of Different Sentiment Lexica for Sentiment Analysis of Tweets. In: *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2015)*. Hissar, Bulgaria <https://www.aclweb.org/anthology/R/R15/R15-1064.pdf>
- Pian A. (2012). Un terrain dit «sensible» dans le champ des migrations : réflexivité sur une expérience marocaine, *E-migrinter*, n°9, p. 79-90, http://www.mshs.univ-poitiers.fr/migrinter/e-migrinter/201209/emigrinter2012_09_079.pdf