Contents lists available at ScienceDirect

Biological Conservation

Content analysis of newspaper coverage of wolf recolonization in France using structural topic modeling

Marie Chandelier[a,b,*], Agnès Steuckardt[b], Raphaël Mathevet[a,c], Sascha Diwersy[b], Olivier Gimenez[a]

[a] CEFE, Univ Paul Valéry Montpellier 3, CNRS, Univ Montpellier, EPHE, IRD, Montpellier, France
[b] PRAXILING UMR 5267, Univ Paul Valéry Montpellier 3, CNRS, F34000, Montpellier, France
[c] French Institute of Pondicherry, UMIFRE21 CNRS/MAEE, Puducherry, India

ARTICLE INFO

ABSTRACT

Populations of large carnivores are recovering in Europe and incur increasing conflict interactions with human activities. According to the agenda-setting theory, the media dissemination of information on these interactions is likely to contribute to shaping public perceptions of large carnivores. We conducted a content analysis of printed media coverage of wolf recovery in France over the period 1993–2014, ever since its natural return to southeast France. To do so, we used a recently developed statistical method – structural topic modeling – that allows to generate topics from large amount of texts and formulate new or assess existing hypotheses. This method formally includes covariates to explain variation in topic prevalence and content in a way that is similar to standard regression analyses. We contrasted content variation between articles in a regional (*Nice-Matin*; N = 742) and a national (*Le Monde*; N = 148) newspaper and analyzed time trends in topic prevalence. The most represented topics were mainly related to the management issues regarding wolf recovery. We found that *Le Monde* represented management issues in a generic manner associated with a perspective centred on carnivore species. In contrast, articles in *Nice-Matin* were about factual issues and associated with a human-centred viewpoint. This contrasted framing emphasizes the gap in representations of wolf management between citizens who directly interact with the wolf and favor detailed information content, centred on human views, and citizens who do not interact or only indirectly with the wolf who will focus on less detailed news, with a more ecological approach. We suggest that increased communication between local and national stakeholders and institutions could provide the context for a more balanced media content of interactions between carnivore species and human activities. This combination could attenuate the gap between regional and national representations.

1. Introduction

Interactions between humans and nature increasingly lead to human-wildlife conflicts (HWCs) (Redpath et al., 2013). With many stakeholders involved, HWCs occupy a substantial place in the establishment of environmental policy and are widely broadcasted to a large audience through mass media. In this context, the media coverage of environmental issues is being studied to better understand the social representations relayed by the media and their potential impacts on citizens (Gore and Knuth, 2009; Jacobson et al., 2012; Bhatia et al., 2013; Sakurai et al., 2013).

According to the agenda setting theory, the media is likely to have an impact on the representation of an issue, not necessarily by telling citizens *how* to think (opinions) but more likely by showing them *what*

to think, i.e. what should be the main themes about a specific question (McCombs and Shaw, 1972; McCombs, 2005). In this context, learning about the themes relayed in the media is a critical issue for conservation scientists who need to understand and feed social communication (Moscovici, 1977).

The interest of ecologists and conservation biologists for the study of the coverage of environmental issues is increasing. For example, the representation of climate change and global warming has been studied in a diachronic perspective through content analysis (Billett, 2010; Brossard et al., 2009), a method that seeks to extract information from texts using descriptive statistics (Krippendorff, 2004). Content analysis has also been used to study HWCs, most often involving animals that potentially may harm people (Jacobson et al., 2012; Bhatia et al., 2013; Sakurai et al., 2013). Usually, content analysis of an environmental

---

issue consists in the identification of the main themes (topics) and whether the coverage is positive or negative (valence) over a specific time period. By analyzing the topics, the valence of the articles (Jacobson et al., 2012; Bhatia et al., 2013) and the quoted sources (Brossard et al., 2009; Jacobson et al., 2012), content analyses allow a better understanding of how environmental issues are covered by media. With these methods, themes are defined by coders and are therefore dependent on their interpretations from the very beginning of the analysis.

In this study, our objective is to demonstrate that the framework of topic modeling (Blei, 2012), a statistical method to analyse topics and their content, enables to generate topics independently from the coders' intuition, and formulate new or assess existing hypotheses by formally investigate temporal trends and explain variation in content, in a way that is similar to standard regression analyses. Of particular interest, structural topic modeling (STM; Roberts et al., 2013) allows assessing the relationship between topic (content and proportion) and explanatory covariates such as author, date or episodic vs. thematic articles. By studying trends in topic proportions over time and explaining content variations with external factors, STM allows taking the context into account for a better understanding of the dynamic of social representations in a comparative and diachronic view. STM analyses have been developed and mostly used in political (Bauer et al., 2016; Grimmer and Stewart, 2013; Munksgaard and Demant, 2016) and education (Reich et al., 2015) sciences. It is only recently that the method has been applied in environmental sciences with a study showing how important topics can be extracted from thousands of scientific articles (Westgate et al., 2015) and to explore representations of climate change (Lynam, 2016).

We use STM in a case study of HWC. To this aim, the extensive cover by national and local press of the return of wolves to France since 1992 has been particularly interesting in terms of the wide range of stakeholders strong positions that have been revealed, even though some are not even directly involved in the conflict. Once considered as one of the most widespread species in the northern hemisphere, the wolf was extirpated from a large part of its habitats by humans during the last few centuries (Mech and Boitani, 2003). In North America and Europe its contemporary recovery is due to multiple factors including a radical change of the predator's perception (Chapron et al., 2014). From being hunted for reward because of its destructive impact on human activities, especially livestock breeding, the wolf has now become a strictly protected species under the international law, and is considered as a keystone species embodying the symbol of wilderness (Caro, 2010). In this context, the interaction of wolves with agricultural activities in a landscape traditionally devoted to sheep grazing has led to strong conflicts between stakeholders arguing for wolf protection and those demanding the downgrade of its legal status, if not the complete removal of the species. In France, the wolf has spontaneously recolonized from Italy, first in southeastern Alps (Mercantour national park), before colonizing a large part of the French mountain areas (Louvrier et al., 2017). Since 2004, wolf management has been based on a derogation to the Habitat's Directive that allows for a wolf removal quota and the establishment of regulation policy (Duchamp et al., 2017). This strategy has placed the French government in an equivocal situation whereby the strictly protected type of management is associated with an increasing wolf removal. Although this dual management was first elaborated to mitigate the conflict, farmers are still impacted by predation and pursue their claims to downgrade the wolf protection status, whereas ecologists remain opposed to the current regulation policy (Lalo and Degeorges, 2017). The confrontation of those claims is frequently associated with an argumentation either based on an ethic centred on the inclusion of animals, plants and land to the community of life (Leopold, 1949) or on a worldview where humans are the primary holders of moral standing (Lombardi, 1983). It often leads to the exclusion of either the wolf or the breeder, feeding the roughness of the conflict instead of its mitigation (Chandelier et al., 2016).

The media began covering this HWC soon after the first wolves were detected in the south-east of the Maritime Alps in 1992. Newspapers in particular built a specific representation of the debates and the events related to wolf presence, and thus took a position regarding the nature- or the human-centred main approaches. In this paper, our aim is to identify whether national and regional newspapers are closer to a nature-centred or human-centred worldview. We collected a corpus of French printed press articles covering the period from 1993 to 2014. We analyzed variations between the framing of a HWC statewide and in the area where the wolf has recolonised (the Alps), to better apprehend social representations of the conflicts in the media at different geographic scales. We analyzed temporal trends in the topics highly represented in the corpus and explained vocabulary variation for each topic with type of newspaper (regional vs. national) as a covariate. We hypothesized that the geographical diffusion and the target audience of the media would impact topic prevalence. Specifically, because the regional newspaper *Nice-Matin* is published and disseminated in the Alpes-Maritimes – the area where the species is most present and livestock farmers most impacted – it is more likely to emphasize human aspects of the wolf presence. On the contrary, the state-wide diffusion of *Le Monde* may induce a more general perspective of the issue, based on large carnivore protection, and thus be closer to a holistic and nature-centred approach.

2. Methods

2.1. Data collection and pre-processing

The data collected for content analysis are printed press articles published in local and national French daily newspapers over the period 1993–2014. We considered a local press group *Nice-Matin* and a leading national newspaper *Le Monde*. The press group *Nice-Matin* covers the south eastern area of France, and in particular the Alps where the wolf was first re-observed in 1992 and has been present since then. We collected articles following the relevance sampling search process (Krippendorff, 2004). We selected documents containing the word 'loup' ('wolf') from the 'Europresse' database (http://www.europresse.com/default.aspx) for *Le Monde*, and from local archive center for *Nice-Matin*. We then removed articles in which the word 'wolf' did not refer to the animal.

In another step, we selected documents that were exclusively dedicated to the wolf, hereafter *primary* articles. In order to identify such articles, we considered a representative sample by selecting documents with the word 'loup' ('wolf') in the title. Based on this sample, we calculated the ratio of the number of occurrences of the word 'loup' ('wolf') over the total number of words in each document for *Nice-Matin* and *Le Monde* separately, then for each newspaper we identified the boundaries of the interval that contained 95% of the sample distribution of ratios. A threshold was defined as the lower boundary of this interval; above that threshold, an article was qualified as primary.

The articles were then preprocessed by word stemming to keep only the root of a word and the removal of punctuation and stop words that contributed little to topic determination. These operations were conducted with the tm package (Feinerer et al., 2008) in R (R Core Development Team, 2015). Hereafter, we used "word" to refer to truncated words that result from the word stemming.

2.2. Statistical analyses

We used topic modeling to interpret patterns in our corpus through a small number of thematic elements – the topics – that were statistically defined using sets of words that co-occurred with unusual frequency (Blei et al., 2003). We used STM that has recently been developed to specifically incorporate covariates measured at the level of the document, here the newspaper articles (Roberts et al., 2016). These covariates can be incorporated to explain variation in *prevalence*, i.e. how much of an article is