

**NAMING DISPLACED PEOPLE: NEW PATTERNS IN
MEDIA DISCOURSE?
A DISCOURSE ANALYSIS OF *LE MONDE* AND
*LE FIGARO***

NOMEANDO AS PESSOAS DESLOCADAS:
NOVOS PADRÕES NO DISCURSO MEDIÁTICO?
UMA ANÁLISE DO DISCURSO EM *LE MONDE* E *LE FIGARO*

Laura Calabrese & Valérie Mistiaen*

vmistiae@ulb.ac.be

lcalabre@ulb.ac.be

The *migrant crisis* has received huge media coverage and has been the subject of many social controversies, among which the one sparked by *Al Jazeera* regarding the words *migrant* and *refugee* during the summer 2015. This study addresses the question of whether the lexical debate that followed had a permanent impact in journalistic writing patterns, by analysing media typifications of displaced people. Using a mixed methodology of corpus linguistics and discourse analysis, we analysed a corpus of 376,217 words from the two main French broadsheet newspapers, *Le Monde* and *Le Figaro*, in order to observe if the lexical debate influenced the choice of words of the journalists. The results of the study show that if some changes are visible after the debate (a higher frequency of *refugee* as well as an accurate usage of this legal term), the latter did not prevent journalists from using the word *migrant* (in spite of the negative connotations *Al Jazeera* decried). The study concludes with some hypotheses about the future of those terms, as they will continue expanding their meaning and their referent according to historical events.

Keywords: migrants, refugees, lexical controversy, media discourse, typifications.

A *crise dos migrantes*, massivamente abordada pela comunicação social, encontra-se no centro de numerosas controversas entre as quais a discussão encetada pelo canal *Al Jazeera* a propósito do uso dos termos *migrante* e *refugiado* durante o verão 2015. Tomando como exemplo os termos usados para designar as pessoas em deslocação, esta investigação procura saber se a referida discussão teve um impacto permanente sobre as práticas de escrita jornalística. Para responder a esta pergunta, analisaremos um corpus de artigos de jornais franceses de referências, *Le Monde* e

* Université libre de Bruxelles, Belgique.

Le Figaro, segundo as perspectivas da análise do discurso e da linguística de corpus. Os resultados demonstram, por um lado, uma transformação visível depois da discussão (nota-se uma maior frequência da utilização da palavra *refugiado*, mas também um uso legal adequado do termo). Por outro lado, a análise de corpus revela que a palavra *migrante* continua a ser usada pelos jornalistas, a pesar das conotações negativas denunciadas por *Al Jazeera*, por vezes em situação de sinonímia. A conclusão do artigo apresenta hipóteses sobre o futuro destes termos, especializados na designação das pessoas em deslocação, cuja significação vai continuar a desenvolver-se em função dos referentes e dos acontecimentos históricos.

Palavras-chave: migrantes, refugiados, controversa lexical, discurso mediático, categorização.



Introduction

The year 2015 was a breaking point in the media representation of displaced people. Several events, such as the war in Syria, shipwrecks in the Mediterranean, the arrival of people fleeing war from the Middle East as well as the different reactions of European countries have been summarised under the denomination *migrant crisis*. This event denomination, widely spread in Western media, has been challenged by a few voices, claiming that there is no migrant crisis but a humanitarian crisis. However, there was little discussion about the identity of the people on the move: they were migrants seeking for asylum in Europe. Unless until three years ago, when the Qatari outlet Al-Jazeera English (AJ) decided to stop using the word *migrant* and systematically shift to *refugee*, sparking what has been called a 'debate' in international media. As the editorial decision of AJ forced Western media to reconsider the predominant lexicon of the media coverage and raised awareness about the usage of actors' denominations, we would like to explore if it provoked a permanent shift in the usual representation of migrants.

The image of migrants, immigrants, refugees and asylum seekers in media discourses has been widely studied in Discourse Analysis, mainly from the perspective of social representations of minorities and usually with a critical stance, aiming at denouncing a bias in media discourse (Van Dijk, 2006). Among this research, some has focused on lexical issues, as it has become obvious for linguists that word choices can change the whole

interpretation of a historical event. In a study conducted on British media, Gabrielatos and Baker (2008) showed that the different co-texts surrounding *refugee* and *asylum seeker* can be explained by the fact that newspapers operate according to the dictionary definition, and not the legal one. However, they found that all the keywords under scrutiny overlap at some point, which is “indicative of the ‘misuse of terminology’ [...] and the interchangeable use” (Gabrielatos & Baker, 2008, p. 26) of *refugee/asylum seeker* with *immigrants/migrants*. Another interesting fact from their corpus, spanning a period from 1996 to 2005, is that *migrants* had a lower frequency than *refugees*, *asylum seekers*, and *immigrants*, which shows that even if the word is available in the English lexicon, it did not correspond to a prominent social phenomenon at the time of the study. Working on the same corpus, Baker, Gabrielatos, Khosravinik, Krzyzanowski, McEnery and Wodak (2008) observed an overlap between the terms used to name displaced people in the British press. By examining the shared consistent collocates¹ of each term, they observed that *refugees*, *asylum seekers*, *immigrants*, and *migrants* are used as near-synonyms. From another perspective, Holmes and Castañeda (2016) showed how German media demarcate the ‘deserving’ *refugee* from the ‘undeserving’ *migrant*, as a consequence of new political discourse patterns that restricts migration in the European Union. The abovementioned studies show that focusing on the lexicon used by journalists allows us to observe the emergence of new categorisation patterns in media discourse.

Lexical and semantic social debates are worth being studied because social phenomena are mainly made of discourse. Indeed, denominations always carry a viewpoint (Siblot, 2001), and are often a battlefield for social actors who try to impose their own meaning or agenda (Krieg-Planque, 2009). Moreover, denominations are the main condition for social phenomena to exist, as they rely on discourse and not on material reality (Kaufmann, 2006; Searle, 1995). If a wide repertoire of lexical and syntactic resources is available for social actors, they will make choices according to their social and ideological position, which will generate discourse patterns (see for instance Fang, 1994). Eventually, those choices will be made by default or dictated by a professional routine, as it is often the case in journalistic writing (Palmer, 2006). If journalists are usually aware of their lexical choices, most of the time they need to use the language in a referential,

¹ Collocation is defined as all the words that frequently concur with a node in a specified span (Sinclair, 1991, p. 105). Consistent collocates refer to collocates that are not seasonal or context-dependent.

not metalinguistic, fashion. This is precisely why the debates around words to describe reality are a perfect observatory for linguists, as they reveal a breach in the social world. When journalists discuss the meaning and usage of the words *terrorist*², *ethnic cleansing*³, *islamophobia*⁴ or *undocumented migrants*, to mention just a few, they try to calibrate words in order to report accurately a reality that is objective to them, thus showing that the interface we use to refer to a common world (language) is not perfect.

The lexical debate that took place a few years back regarding the categorisation of displaced people is very symptomatic of the socio-political tensions in a globalised world, as it reveals the will to build social reality in a certain way. In a previous research conducted on the debate sparked by AJ's article (Calabrese, 2018), we found that several media outlets reacted to it, publicizing the decision and creating the context for a lexical discussion. A corpus of 22 newspaper articles (in English, French and Spanish) were analysed using a qualitative methodology. All the articles addressed the question of the lexical and semantic problems raised by the words of displaced people, and featured lots of metalinguistic terms, among which the expression 'semantic debate' was predominant. Following that research, we wanted to know if the debate had a permanent impact on journalistic routines and therefore on the denomination patterns used to name displaced people. In order to achieve that goal, this paper will utilise a much larger dataset. A corpus of articles from the two main French broadsheet newspapers with the largest circulation (*Le Monde* and *Le Figaro*) has been compiled and analysed using a mixed methodology of Discourse Analysis (DA) and Corpus Linguistics (CL), in order to validate the following hypotheses:

H1. Following AJ's editorial decision, and as it became clear that it was loaded with negative connotations, the word *migrant* became less frequent in daily French broadsheet newspapers;

H2. On the other hand, as the debate raised awareness on the usage of words and their legal implications, the word *refugee* stopped being used as a co-referent (e. g. referring to the same social phenomenon) of *migrant*, but

2 "BBC. Editorial Guidelines": <<http://www.bbc.co.uk/editorialguidelines/guidance/terrorism-language/guidance-full>>

3 "'Homophobia' and 'Islamophobia' are the right words for the job" (*The Guardian*, 27/11/2012) <<https://www.theguardian.com/commentisfree/2012/nov/27/homophobia-islamophobia-right-words-associated-press>>

4 "Islamophobie: un abus de langage" (*Libération*, 20/9/2013) <http://www liberation.fr/france/2013/09/20/islamophobie-un-abus-de-langage_933600>

only in relation with its legal meaning (someone who obtained the asylum status).⁵

In the case that H1 and H2 would prove right, the corpus will present no overlap between the two terms.

Before tackling the corpus analysis, we will explain the context of AJ's article and the reason why it became such a successful media event. In the second part, the corpus and methodology will be laid out, and in the third part the results of the corpus analysis will unfold.

The context of the debate

On the 20th of August 2015, Barry Malone, online editor of the English version of Al-Jazeera, published a blog post announcing that the media outlet would no longer use the word *migrant* to cover current events:

The umbrella term *migrant* is no longer fit for purpose when it comes to describing the horror unfolding in the Mediterranean. It has evolved from its dictionary definitions into a tool that dehumanises and distances, a blunt pejorative.⁶

The article can be seen as a discursive event, namely a speech-related media event caused by the particular position of the speaker (see Calabrese, forthcoming). Indeed, several Western media reacted to it with by publishing metalinguistic articles, creating the illusion of a pre-existing debate⁷ on lexical issues. Discursive events are usually provoked by politicians, whose words are publicised and promoted by journalist. They are considered vital for the former, who are fond of publicity, as for the latter, who need to cover current events. AJ's lexical decision, if not unusual for a media outlet, goes far beyond the usual behaviour of media discourse, as it was seen as "an interventionist challenge to Western media representations of immigration

5 We prefer to say *referent* instead of *synonym* because before the debate, we do not know if journalists considered both words as synonyms. On the other hand, given the discursive usages, we can assume that they considered them as interchangeable in most contexts and referring to the same phenomenon.

6 "Why Al Jazeera will not say Mediterranean 'migrants'" <<http://www.aljazeera.com/blogs/editors-blog/2015/08/al-jazeera-mediterranean-migrants-150820082226309.html>>

7 It is worth mentioning that the so-called debate was mostly an outcome of the international coverage of AJ's article.

[...] a political move which usurped European sovereignty” (Kyriakides, 2016, p. 2). By reframing the event (shipwrecks, migrants’ deaths in the Mediterranean, asylum applications and border control) around a lexical issue, AJ takes an explicit stance in favour of welcoming extra-European migrants, bringing pressure to bear on a very restrictive asylum policy (Valluy, 2005), as it can be read in this article published by *The Huffington Post*:

Tim Stanley, historian and Daily Telegraph columnist, told the programme he was ‘sympathetic’ to journalists who try to be sensitive in their coverage of the crisis. But he said, though the vast majority of them were refugees according to the UN, calling all of them this word was an attempt to “politically put the onus on Europe to accept everyone without conditions and without due process.” (“Al Jazeera Denies ‘Politicising’ Migrant Crisis In Deciding To Call People ‘Refugees’ Instead”, 25/08/2015)

It must be said that the social context was profitable to AJ’s article. First of all, the dramatic events taking place in the Mediterranean since 2013 were (and still are) pervasive in the news media. Second, the lexical and discursive “profile” (Veniard, 2013) of the words related to migration were (and still are) clearly negative, with a news coverage based on the metaphor of the “flood” (Gabrielatos & Baker, 2008; Kosnick, 2014). Third, AJ’s article comes forth at a time where the hierarchy between people on the move is already discussed in news media as well as in social media, even though the discussion is scarce and weakly organised.⁸ Indeed, before the debate sparked by AJ’s post, some articles tackled the problem of social categorisation:

(1) Why are white people expats when the rest of us are immigrants? (the-guardian.com, 13/03/2015)

In the lexicon of human migration there are still hierarchical words, created with the purpose of putting white people above everyone else. One of those remnants is the word “expat”.

It can be seen that the context contributed to the success of the discursive event in several ways. First of all, the massive circulation of images of humanitarian catastrophes, the negative connotation of certain words (but

⁸ These hierarchies can be summarised with the well-known Zygmunt Bauman’s categories of the tourist and the vagabond: “Green light for the tourists, red light for the vagabonds” (Bauman, 1998, p. 93).

not others, e.g. *expat*) related to migrations, and finally a metalinguistic reflection surrounding those words followed by the acknowledgement of an existing hierarchy between them. But in spite of the impact of AJ's strong statement, the use of words related to migrations are legally bound, as the status of refugee implies several requirements⁹ that cannot be overlooked when referring to displaced people.

Corpus and Methodology

The main corpus comprises 339 articles (282,478 words) from the paper versions of *Le Monde* and *Le Figaro*, the two national French daily broadsheets with the largest audience. As the latter has a conservative and the former a central-left agenda, we can expect that their choice of language reflect their political stance:

It does not seem controversial to suggest that the choice of words to be used in relation with RASIM [Refugees, asylums seekers, immigrants, and migrants] can be used as a clear indication of the stance of the writer/newspaper toward these groups-particularly when the phraseology used is either compatible to, or unwarranted by, the definitions of these terms [...]. (Gabrielatos & Baker, 2008, p. 14)

The articles were collected during some specific periods where the words under scrutiny were expected to appear the most. Each period was chosen according to current events related to migrations to/through Europe. Such periodisation would allow to observe seasonal collocates and eventual shifts in the usage of words. The following periods were studied:

1. 7th-20th September 2015: two weeks after AJ's article.
 - a. 7th-13th September 2015
 - b. 14th-20th September 2015
2. 1st-7th January 2016: Cologne sexual assaults.

9 The Geneva Convention from 1951 stipulates that a refugee, "(2) As a result of events occurring before 1 January 1951 and owing to well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group or political opinion, is outside the country of his nationality and is unable or, owing to such fear, is unwilling to avail himself of the protection of that country; or who, not having a nationality and being outside the country of his former habitual residence as a result of such events, is unable or, owing to such fear, is unwilling to return to it" (<http://www.ohchr.org/EN/ProfessionalInterest/Pages/StatusOfRefugees.aspx>).

3. 25th-31st May 2016: on the 29th of May, it is thought that 700 people have died in a shipwreck in the Mediterranean Sea.
4. 17th-24th October 2016: Calais Jungle demolition.
5. 8th-14th May 2017: week of shipwrecks in the Mediterranean Sea.
6. 3rd-9th of July 2017: Joint declaration on migration by EU Commissioner Avramopoulos and the Ministers of Interior of France, Germany and Italy.

In order to check the validity of the results, a smaller preliminary corpus composed of 122 articles (93,739 words) was created, referring to two different periods:

1. Just before AJ's article:
 - a. from the 6th to the 12th of August 2015
 - b. from the 13th to the 19th of August 2015
2. Right after AJ's article:
 - from the 20th of August to the 28th of August 2015
 - from the 29th to the 6th of September 2015

For each period, all the articles mentioning one of the two words were collected, by means of a lexical query in the Europresse¹⁰ database of *réfugié** 'refugee' and *migrant** 'migrant', including the plural and female nouns. From the list obtained following the query, we removed articles from press agencies, articles with less than 300 words, articles mentioning only the word *réfugiés* in the proper name UNHCR, and articles that did not have the current *migrant crisis* as the main topic.

To carry out the analysis, two partitions of the corpus were tried: by newspaper (see Figure 1) and by period (see Figure 2). In the second case, the corpus was subdivided into three subcorpora, in order to observe if the two lexis were used indifferently, as it was the case before the lexical debate began:

- Articles mentioning both *réfugié** and *migrant**;
- Articles mentioning only *réfugié**;
- Articles mentioning only *migrant**.

¹⁰ Europresse: <<http://www.europresse.com/fr/>>

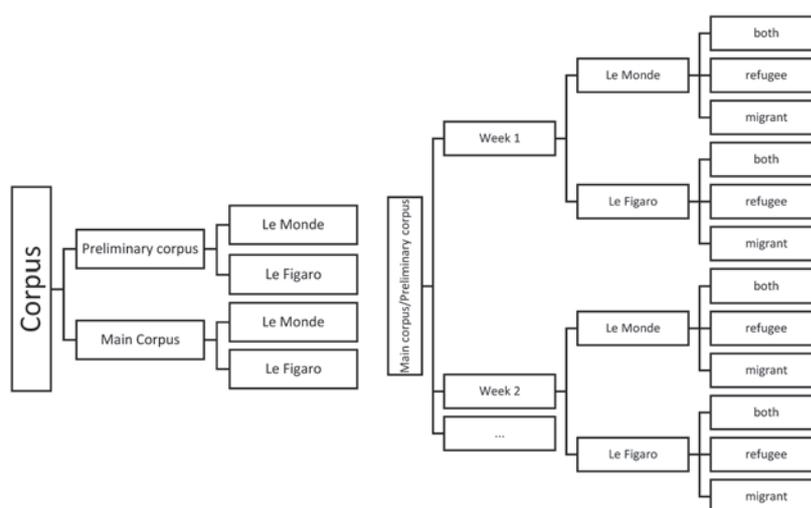


Figure 1. Corpus representation by newspaper

Figure 2. Corpus representation by period

According to H2, after the debate there will be no overlapping of the words, as they will no longer be considered as co-referents. As the preliminary corpus (PC) is supposed to reflect the general use of both lexis before AJ's statement, by comparing it with the main corpus (MC), it will be possible to check if there were more articles using both terms as co-referents before than after the debate.

For this study, the methodology is a mix of corpus linguistics (CL) and discourse analysis (DA).¹¹ If DA provides the concepts and theoretical framework to analyse discourse,

Using a "Corpus linguistics methodology allows for a higher degree of objectivity—that is, it enables the researcher to approach the texts (relatively) free from any preconceived notions regarding their linguistic or semantic/pragmatic content. When the starting point is keyword analysis, the analyst is presented with a list of words/clusters which will then be examined in (expanded)

¹¹ Under the label DA we consider French Discourse Analysis as well as Critical Discourse Analysis, two theoretical frameworks belonging to the same family as they share a common origin in Foucault's and Bakhtin's writings. If they developed a different set of concepts, they remain perfectly compatible and particularly complementary.

concordances for their patterning and contextual use” (Gabrielatos & Baker, 2008, p. 7).

After gathering all the PDF articles, they were processed by a corpus analysis toolkit for concordance and text analysis, *AntConc*.¹² The following features were employed:

- 1) The keyness determines the keywords of the corpus, by pointing out the most frequent lemmas. The results show that *réfugié** and *migrant** are the main keywords of the global corpus.
- 2) The frequency of both terms (*réfugié** and *migrant**) diachronically determines which one is more often used than the other, and when changes occur in time (frequency by period of time). This feature will allow us to verify H1.
- 3) The concordance tool analyses the collocates of a term, which contribute to its meaning, and their examination can provide “a semantic analysis of a word” (Sinclair, 1991, p. 116). The most frequent collocates were sorted by categories (cf. *infra*) and the frequency of these words were then counted in association with both lemmas. This concordance analysis was repeated for each period of time in order to study the corpus through a diachronic lens. The span was set at five words to the left and right of the node. This feature will allow us to observe if the two lexis are used in very different co-texts, or on the contrary, as co-referents, and thus validate or invalidate H2.

The tools that CL offers can help answering traditional questions raised by DA, especially under which social conditions a lexical shift occurs. Drawing on this tradition, our research aims at explaining “why and under what circumstances and consequences the producers of the text have made specific linguistic choices among several other options that a given language may provide” (Baker *et al.*, 2008).

Results

The results of the analysis will be presented under the form of graphs and tables in order to visualise the main patterns of the corpora, and then commented in order to prove H1 and H2.

¹² Laurence Antony's AntConc software: <<http://www.laurenceanthony.net/software.html>>

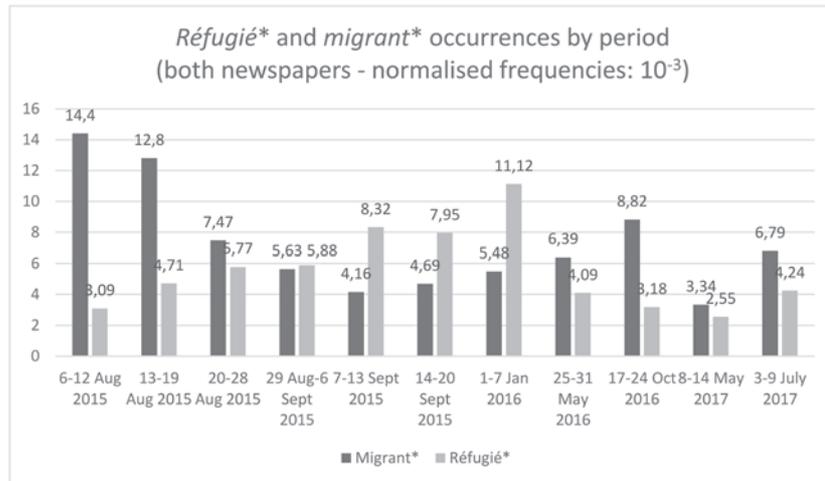
Frequencies of the lemmas before/after the debate

The wordlist generated a list of the most frequent words for each corpus. Table 1 shows the first ten results:

Preliminary corpus (Number of words: 93,739)			Main corpus (Number of words: 282,478)		
Rank	Frequency	Lemma	Rank	Frequency	Lemma
21	570	migrants	27	1447	réfugiés
30	414	réfugiés	36	1013	migrants
34	387	pays 'land'	38	939	accueil 'welcome'
39	325	Europe	41	870	pays 'land'
42	282	asile 'asylum'	48	693	France
45	261	accueil 'welcome'	55	599	politique 'politics'
49	231	frontières 'borders'	56	590	asile 'asylum'
59	175	politique 'politics'	59	566	Europe
60	166	Allemagne 'Germany'	61	548	Allemagne 'Germany'
74	129	crise 'crisis'	62	542	frontières 'borders'

Table 1. First ten words of the preliminary and the main corpus wordlists (raw frequency by corpus, *Le Monde* and *Le Figaro*)

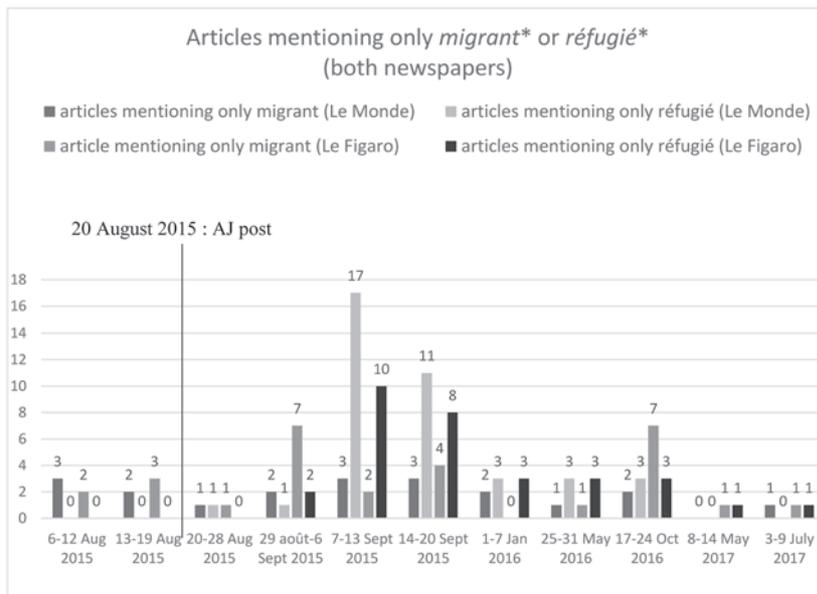
This table shows that in the PC, *migrant** scores higher than *réfugié**, whereas in the MC the opposite happens. Even though the PC is smaller than the MC, the rank both words respectively hold in each corpus is quite eloquent. It is clear that a shift in the use of the terms has occurred. Nevertheless, this observation does not suffice to prove H1. If we take a look at the frequencies per period, it is possible to see a breaking point that can be situated in September 2015:



Graph 1. *Réfugié** and *migrant** occurrences by period, normalised frequencies 10^{-3} (*Le Monde* and *Le Figaro*)

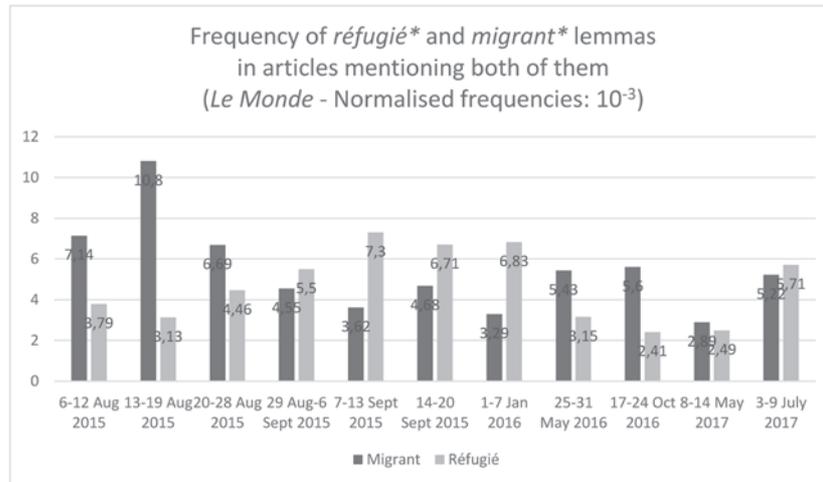
Graph 1 features a clear pattern, with more occurrences of *migrant* during August and the beginning of September 2015, the reverse trend during September 2015, and a return to the first pattern at least from May 2016 on, more than one year after the debate, with *migrant* taking again the advantage over *réfugié*.

Graph 2 (articles that mention either *réfugié* or *migrant*, but not both) supports this observation, by showing that before the debate, there are barely any articles mentioning only *réfugié*, but on the other hand a spike occurs during September. It is interesting to note that even if *Le Figaro* features the same trend as *Le Monde*, due to its conservative political agenda it shows a preference for *migrant* right before the spike.

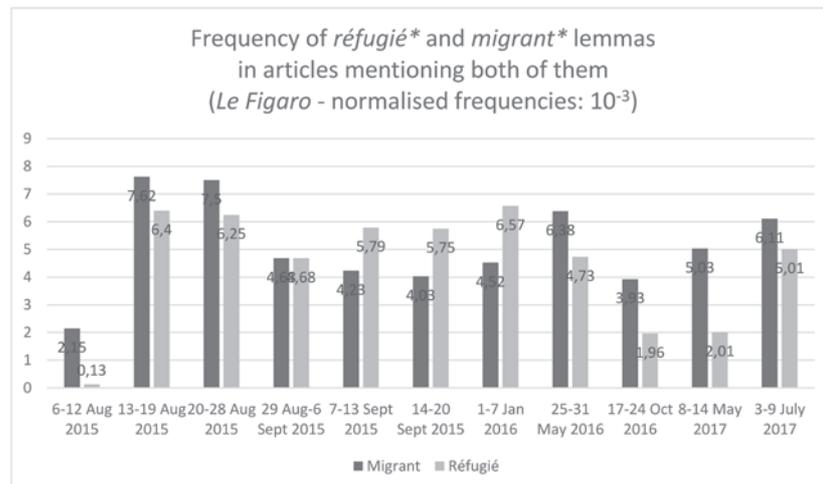


Graph 2. Articles mentioning *réfugié** but not *migrant** or *migrant** but not *réfugié**, raw frequencies (*Le Monde* and *Le Figaro*)

If we look carefully at the articles mentioning both terms by period (Graphs 3 and 4), it can be noticed that *Le Monde* and *Le Figaro* follow a similar pattern, with more occurrences of *migrant* during August and the beginning of September 2015, the reverse trend during September 2015 (with a spike of *réfugié*), and a return to the first pattern at least from May 2016 on.



Graph 3. Normalised frequencies (10^{-3}) of *réfugié** and *migrant** lemmas in articles mentioning both of them (*Le Monde*)



Graph 4. Normalised frequency (10^{-3}) of *réfugié** and *migrant** in articles mentioning both lemmas (*Le Figaro*)

Once again, the political agendas of each newspaper could explain a slight divergence in the frequencies: if the shift is immediately visible in *Le Monde* (165 occurrences of *réfugié* for 142 *migrant* starting the week of the 29/08/15), it is less clear for *Le Figaro*, as both terms are equally mentioned the same week (99 occurrences of both terms).

Having examined both corpora under different angles, it became clear that an occasional shift occurred right after the debate, only to come back to the initial journalistic writing routine. This being said, the CL analysis does not show all the data. A thorough observation of the corpus reveals a more consistent change in the headings, namely the umbrella denomination that journalists use to create event families. In *Le Monde*'s PC, we only find the heading *crise des migrants* 'migrant crisis', but a few weeks later, in the MC, there are 16 occurrences of *crise des réfugiés* 'refugee crisis' for only 9 occurrences of *crise des migrants* (and 1 *crise migratoire* 'migratory crisis'). As if *Le Monde* wanted to go further, it even named a section *repenser la crise des réfugiés* 'rethinking the refugee crisis', 5 occurrences).

<i>Corpus</i>	<i>Crise des migrants</i>	<i>Crise des réfugiés</i>	<i>Crise migratoire</i>	<i>Repenser la crise des réfugiés</i>
Preliminary corpus	8 15,38	0 0	0 0	0 0
Main corpus	9 5,26	16 9,36	1 0,58	5 2,92

Table 2. Headings, normalised frequencies 10^{-2} in italics (*Le Monde*)

Le Figaro does not present such heading names, as we found only one *crise des migrants* in the MC, but *immigration* is used 7 times in the PC and 9 times in the MC.

If we examine the event names, namely the collocation used by the journalists to name this specific *crisis*, the paradigm unfolds as follows:

Corpus	<i>Crise des migrants</i> + <i>Crise de migrants</i>	<i>Crise migratoire</i>	<i>Crise des réfugiés</i> + <i>Crise de réfugiés</i>
Preliminary corpus	(9 + 0) 9 19,04	5 10,58	(2 + 1) 3 6,35
Main corpus	(15 + 0) 15 9,28	30 18,57	(15 + 0) 15 9,28

Table 3. Event denominations excluding the headings, normalised frequencies (10^{-5}) in italics (*Le Monde*)

Corpus	<i>Crise des migrants</i> + <i>Crise de migrants</i>	<i>Crise migratoire</i>	<i>Crise des réfugiés</i> + <i>Crise de réfugiés</i>
Preliminary corpus	(9 + 2) 11 23,67	8 17,21	(6 + 1) 7 15,06
Main corpus	(22 + 1) 23 19,02	18 14,89	(8 + 0) 8 6,62

Table 4. Event denominations excluding the headings, normalised frequencies (10^{-5}) in italics (*Le Figaro*)

If all denominations are co-referents in media discourse, the signifier cannot be ignored, as it shapes our perception of the event (Calabrese, 2013). Even if *crise des migrants* ‘migrant crisis’ and *crise migratoire* ‘migratory crisis’ belong to the same family, we must distinguish both collocations, as the role of the noun and that of the adjective are paramount in the way we perceive a social crisis. Regarding the syntax, the format *event name + adjective* (*crise migratoire* ‘migratory crisis’) gives a more synthetic view of the event (as there is a continuity between the two elements) than the format *event name + article + noun* (*crise des migrants* ‘migrant crisis’). Moreover, the adjective modifies the noun *crisis*, when the noun only specifies the kind of crisis in question. If we consider the lexicon, the second denomination has a focus on the actors, when the first one emphasises the action of migrating. *Crise migratoire* is then a better candidate to avoid the lexical problems raised both by *migrant* and *réfugié*, which can explain the rise in the MC (48 occurrences, for 13 in the PC). Nevertheless, *crise des migrants* still scores better in both corpora.

Headlines have an important cognitive function: they are read first and thus help the reader construct the overall meaning before the text itself is even read, summarise the most important information of the report under the form of keywords, and finally, represent the newspaper's ideological stance (Van Dijk, 1991). The examples below show a particular structure of headlines used by French journalists, with the word left of the colon referring to the "aboutness" (Bosredon & Tamba, 1992) of the text:

- (2) Migrants: le ton se durcit entre la droite et l'exécutif (*Le Figaro*, 11/09/2015)¹³
'Migrants: the tone rises between the right and the government'
- (3) Migrants: des maires français s'engagent (*Le Monde*, 08/09/2015)
'Migrants: French mayors get involved'
- (4) Réfugiés: Sarkozy veut refonder Schengen (*Le Monde*, 11/09/2015)
'Refugees: Sarkozy wants to rethink Schengen'
- (5) Réfugiés: Hollande engage la France dans la logique des quotas (*Le Figaro*, 08/09/2015)
'Refugees: Hollande says France will accept quotas'

These examples clearly show that *migrant* and *réfugié* are not only nouns aiming at categorising individuals according to their status in a territory, but also a media category referring to an event (or series of events) that is supposed to be known to the readers.

Corpus	<i>Migrant*</i>	<i>Réfugié*</i>	<i>Crise des migrants</i>	<i>Crise des réfugiés</i>	<i>Crise migratoire</i>
Preliminary Corpus	21 <i>40,38</i>	6 <i>11,54</i>	0 <i>0</i>	0 <i>0</i>	1 <i>1,92</i>
Main Corpus	31 <i>18,13</i>	28 <i>16,37</i>	0 <i>0</i>	0 <i>0</i>	2 <i>1,17</i>

Table 5. Headlines including *réfugié** or *migrant**, normalised frequencies (10⁻²) in italics (*Le Monde*)

¹³ Our translation of the French text is not literal, but intends to keep the general meaning of the headlines.

Corpus	<i>Migrant*</i>	<i>Réfugié*</i>	<i>Crise des migrants</i>	<i>Crise des réfugiés</i>	<i>Crise migratoire</i>
Preliminary Corpus	23 33,33	7 10,14	0 0	0 0	0 0
Main Corpus	31 18,45	17 10,12	4 2,38	2 1,19	1 0,6

Table 6. Headlines including *réfugié** or *migrant**, normalised frequencies (10^{-2}) in italics (*Le Figaro*)

Again, both Tables 5 and 6 highlight the previously seen trend: *migrant** is less and less used in the headlines of the MC. Interestingly (and contrary to what happens in the articles), *migrant** is still most present than *réfugié** in the headlines of the MC.

So far, we have seen two different trends. On the one hand, the frequencies of *migrant* tend to decrease after the debate in favour of *réfugié* (with a strong spike in September 2015, whether in the whole corpus, in the articles mentioning only *réfugié* or in the articles mentioning both terms), only to come back to the regular pattern after the spike. On the other hand, event names show a more permanent change in journalistic writing patterns. As seen in Tables 2 to 6, this trend is more visible in *Le Monde* than in *Le Figaro*. If the former features a clear-cut rise in *crise de(s) réfugiés* after the debate, which does not occur in *Le Figaro*, both newspapers show a rise in the more neutral form *crise migratoire*.

If the frequencies show that the two terms are still being used by journalists, it is time now to focus on the co-text, meaning and referent of the terms, in order to check if they are used as co-referents, or on the contrary, with a clear distribution of meaning.

Referent and meaning of the lemmas before/after the debate

As pointed out in the introduction, previous research has shown that journalistic discourse tended to use the two terms equally, regardless of the actual legal status of the displaced person. Even if the debate sparked by AJ raised awareness among journalists, the latter could not stick to the term *refugee* as AJ wished, because of the restriction of its legal meaning (see footnote 9); instead, they could avoid the misuse of the terms. The aim of

this section is then to find out whether after the debate the two French newspapers use the terms more accurately, in regard to the legal definition of *refugee*. Indeed, Graphs 2 to 4 show that not only the two French newspapers continued using *migrant**, but they also did it alongside *réfugié**. In order to observe if they are used as co-referents, and validate or invalidate H2, the co-text of the words will be studied by means of the concordance tool. This analysis will be illustrated with samples from the corpus.

On the basis of the most recurring collocates of *migrant** and *réfugié** (detailed in Table 7), a list of categories was created:

- Semantic field of welcoming: *accueil* 'welcome' (noun), *accueillir* 'welcome' (infinitive verb), *accueilli* 'welcomed' (male singular past participle), *accueillis* 'welcomed' (male plural past participle), *accueillie* 'welcomed' (female singular past participle), *accueillies* 'welcomed' (female plural past participle), *accueillent* 'welcome' (third plural person present simple), *accueille* 'welcome' (third singular person present simple).
- Semantic field of stream: *flot* 'flood', *flots* 'floods', *flux* 'flow', *afflux* 'influx', *vague* 'wave', *vagues* 'waves'.
- Event denomination: *crise* 'crisis', *crises* 'crises'.
- Semantic field of confinement: *camp* 'camp', *camps* 'camps', *campement* 'encampment', *campements* 'encampments'.
- Semantic field of quantity: *massif* 'massive', *massifs* 'massive' (male plural), *massive* 'massive' (female singular), *massives* 'massive' (female plural), *millier* 'thousand', *milliers* 'thousands', *centaine* 'hundred', *centaines* 'hundreds'.
- Semantic field of legal status: *guerre* 'war', *guerres* 'wars', *politique* 'political', *politiques* 'political' (plural), *économique* 'economical', *économiques* 'economical' (plural), *quota* 'quota', *quotas* 'quotas'.

Collocates	Réfugié*	Migrant*
<i>accueil*</i>	329	142
<i>flot*</i> , <i>*flux</i> , <i>vague*</i>	72	91
<i>guerre*</i> , <i>politique*</i> , <i>économique*</i>	89	87
<i>réfugié*</i>		77
<i>migrant*</i>	77	
<i>massif*</i> , <i>millier*</i> , <i>centaine*</i>	74	74
<i>crise*</i>	61	66
<i>camp*</i>	40	22
<i>quota*</i>	38	11

Table 7. Collocates of *réfugié** and *migrant** in all corpora (raw frequency)

When analysing the corpora per newspaper, we focused only on the collocates that contribute to the legal meaning of *réfugié*; those that did not (such as *afflux*, *crise*) were removed. In order to establish a list of concepts that contribute to the legal meaning, we considered that refugees (based on the definition of the UNHRC as well as European directives), contrary to migrants, must be welcomed in reception centers (“centre d’accueil”), flee war or other conflicts with a political nature (contrary to an economical one), and are spread through the European territory according to quotas. This means that all those words are expected to be found in the co-text of *réfugié*. The word *camp* is worth being briefly discussed. Even if it is not clearly defined in the legal literature, it belongs to the semantic field of *réfugiés*, as the High Commissioner for Refugees is one of the main actors to handle refugee camps (Valluy, 2005). This could explain the 31 occurrences of *camp(s) de réfugiés* ‘refugee camp(s)’ in the corpus. On the other hand, the word *camp* is available in the common language to form collocations such as *camp de migrants* ‘migrant camp’, with only 5 occurrences in the corpus, alongside an unconventional *campement de migrants* ‘migrant encampment’ (2 occurrences), that could be a hint of the unofficial nature of *migrant camps*.

	<i>Le Monde</i>				<i>Le Figaro</i>			
	Preliminary Corpus (47,261 words)		Main Corpus (161,555 words)		Preliminary Corpus (46,478 words)		Main Corpus (120,923 words)	
Collocation	<i>Mi-grant*</i>	<i>Ré-fugié*</i>	<i>Mi-grant*</i>	<i>Ré-fugié*</i>	<i>Mi-grant*</i>	<i>Ré-fugié*</i>	<i>Mi-grant*</i>	<i>Ré-fugié*</i>
Welcome	17 <i>35.97</i>	35 <i>74.06</i>	55 <i>34.04</i>	161 <i>99.66</i>	10 <i>21.52</i>	7 <i>15.06</i>	49 <i>40.52</i>	108 <i>89.31</i>
War	0 <i>0</i>	2 <i>4.23</i>	2 <i>1.24</i>	20 <i>12.38</i>	3 <i>6.45</i>	2 <i>4.3</i>	1 <i>0.83</i>	20 <i>16.54</i>
Politic	0 <i>0</i>	5 <i>10.58</i>	2 <i>1.24</i>	12 <i>7.43</i>	0 <i>0</i>	5 <i>10.76</i>	0 <i>0</i>	17 <i>14.06</i>
Economic	10 <i>21.18</i>	3 <i>6.35</i>	23 <i>14.24</i>	2 <i>1.24</i>	16 <i>34.42</i>	2 <i>4.3</i>	29 <i>23.98</i>	4 <i>3.31</i>
Quota	2 <i>4.23</i>	7 <i>14.81</i>	4 <i>2.48</i>	13 <i>8.05</i>	2 <i>4.3</i>	7 <i>15.06</i>	4 <i>3.31</i>	12 <i>9.92</i>

Table 8. Collocates of *réfugié** and *migrant**, normalised frequencies in italics (10^{-5}) (*Le Monde* and *Le Figaro*)

According to our results, the concept of welcoming is by far more related to *réfugié* in both corpora, but even more in the MC. The collocation *accueil des réfugiés* ‘reception of refugees’ occurs 105 times, compared to 46 occurrences for *accueil des migrants* ‘reception of migrants’. Even if a few exceptions are reported, both newspapers use most of the time *political refugees*, *war refugees* and *economic migrants*, showing a clear distribution of meaning that matches the dictionary definitions of the terms. *Quota(s)* also occurs more frequently in the co-text of *réfugié*.

With a few exceptions, these results speak in favour of H2, as they seem to confirm that the terms are not interchangeable, but on the contrary, occur in different co-texts. Table 8 shows an evolution between the PC and the MC, in the sense of a specialisation of the meaning of *refugee* after the debate, henceforth surrounded by the collocates that contribute to its legal meaning. But even if this seems to confirm that journalists tried to avoid misnomers, the terms are not systematically used according to the legal definition. An evidence of this are the 77 occurrences of *réfugié* found in the

co-text of *migrant* (or the other way around); even if only some of them are used as co-referents (referring to the same reality) and as synonyms (having the same meaning), they are worth being mentioned:

- (6) Débordée, la Bavière veut accélérer le retour des réfugiés [Title]
Sont concernés les migrants en provenance des Balkans [Subtitle] (*Le Figaro*, 02/09/2015)¹⁴
'Unable to cope with refugees, the Bavarian region wants them to go back'
[Title] 'Migrants coming from the Balkans are mainly concerned by this measure'
[Subtitle]
- (7) Sur la côte turque de la mer Egée, les départs de réfugiés vers les îles grecques ont cessé. Plus un migrant ne passe ("Migrants: L'accord signé entre l'Europe et la Turquie fonctionne", *Le Monde*, 28/04/2016)
'No more departures of refugees to the Greek islands from the Turkish Aegean coast. No more migrants passing'
- (8) La Grèce, "dépassée" par un afflux de réfugiés [Title]
Athènes a promis des mesures pour améliorer l'accueil des migrants, tout en appelant à la solidarité européenne [Subtitle] (*Le Monde*, 08/08/2015)
'Flood of refugees in Greece' [Title]
'Athens promises to improve the reception of migrants and call for European solidarity' [Subtitle]
- (9) [...] près de 800 migrants attendent dans des cabanes la nuit où ils rejoindront l'Angleterre. [...] Tous les soirs, ils déposaient, en fourgonnette, une dizaine de réfugiés sur une aire d'autoroute d'où ils guettaient l'arrivée des camions ("À Calais, les passeurs à plein régime avant la fin de la "jungle", *Le Monde*, 17/10/2016)
'[...] near 800 migrants waiting in shacks to reach England during the night [...] Every night, tens of refugees were left in a motorway, waiting for trucks to pass'

Even if the examples are not numerous, they clearly show that a random usage of the terms in journalistic texts is still possible. At some point, the coexistence of both terms in the same co-text confirms that for several reasons (whether to avoid repetitions, enrich the vocabulary, avoid stereotypes or the difficulty of determining the actual status of displaced people), journalistic discourse still needs to alternate between the two of them.

All the results show that *migrant* will not be out of circulation any time soon. What might happen in the future is that the word will slowly lose some of the negative representations attached to it, as illustrated by some

¹⁴ Our translation of the French headlines is not literal, but allows to observe the overlap of the two terms.

recent examples (not belonging to our corpus) displaying an innovative use of *migrant*, that extends the referent to EU citizens:

- (10) Leaked document reveals UK Brexit plan to deter EU immigrants [Title]
 Exclusive: Home Office paper sets out detailed proposals including measures to drive down number of low-skilled migrants from Europe [Subtitle]
 (*The Guardian*, 5/9/2017).

As events unfold in the current climate of political turmoil in Europe, further research will be needed in order to confirm semantic shifts in the words specialised in naming people on the move.

Discussion

As said in the introduction, there are no real synonyms in groups' categorisations, as each word carries a different viewpoint. Regarding the migrant crisis, several words are available in the common lexicon to refer to people fleeing violence in the Middle East. The lexical debate sparked by AJ had an impact in the representation of displaced people, as it raised awareness of the different meanings carried by each word. One of the limits of AJ's rationale was the legal meaning of the word *refugee*, which refers to a legal status and has obvious consequences in the way an individual is considered by the host State.

The aim of this study was then to observe if the debate had a permanent impact in journalistic writing routines, and therefore in the categorisation patterns they use to name displaced people. In order to do so, we collected, by means of a lexical query in the Europresse database, a corpus of 339 articles (282,478 words) and a preliminary corpus of 122 articles (93,739 words) from the two main French broadsheet newspapers (*Le Monde* and *Le Figaro*). Articles were analysed using a mixed methodology of Discourse Analysis and Corpus Linguistics.

In sight of the results of the corpus analysis, the two working hypotheses are partially validated. Regarding H1 (the word *migrant* is less and less used in daily French broadsheet newspapers), the wordlist shows that in the preliminary corpus of both newspapers, *migrant** scores higher than *réfugié**, whereas in the main corpus the opposite happens. Having examined both corpora under different angles, it became clear that there has been a peak in the use of the word *réfugié* just after the debate, but it has not been permanent, as the previous frequencies of the term rapidly came back.

Nevertheless, *Le Monde* reveals a more consistent shift in the headings. Both newspapers featured a change in the event denomination, with a preference for the more neutral form *crise migratoire* (rather than *crise des migrants* or even *crise des réfugiés*). It is interesting to note that even if *Le Figaro* features the same trend as *Le Monde*, probably due to its conservative political agenda it shows a preference for *migrant* right before the spike. Moreover, it can explain a slight divergence in the frequencies: if the shift is immediately visible in *Le Monde*, it is less clear for *Le Figaro*, as both terms are equally mentioned the week just after AJ's statement.

Regarding H2 (the word *réfugié* will no longer be used as a co-referent of *migrant*), the analysis clearly shows that some collocates are more frequent (if not exclusive) than others, among which the semantic field of welcoming, war, political conflict and the word *quota* hold a preferential place alongside *réfugié*, while *migrant* is more often collocated with *economical*. If these results speak in favour of H2, that is to say, seem to confirm that the terms do not overlap, a few examples showing the use of both words in the same co-text would prove a more unpredictable pattern.

To conclude, even if a few examples showed that the use of both terms is not clearly fixed in journalistic discourse, it seems that the lexical debate did have an impact on the usage of words referring to displaced people, but not exactly in the way the Qatari news outlet expected. If the term *refugee* has been recalibrated around its legal meaning, the word *migrant* has not been removed but rather reinforced in its economical meaning. As current events unfold, further research will be necessary to understand how these two words are encoding new social representations and recording discursive usage as well as semantic change.

References

- Baker, P.; Gabrielatos, C.; Khosravini, M.; Krzyzanowski, M.; McEnery, T. & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse Society*, 19, 273-306.
- Bauman, Z. (1998). *Globalization. The Human Consequences*. Cambridge: Polity Press.
- Bosredon, B. & Tamba, I. (1992). Thème et titre de presse: les formules bisegmentales articulées par un "deux points". *L'information grammaticale*, 54, 36-44.
- Calabrese, L. (2013). *L'événement en discours. Presse et mémoire sociale*. Louvain-la-Neuve: Academia.

- Calabrese, L. (2018). Faut-il dire *migrant* ou *réfugié*? Conflits lexico-sémantiques autour d'un problème public. *Langages*, 210, 105-122.
- Calabrese, L. (forthcoming). L'événement de réception. Un événement de parole du côté des publics. In Brigitte Sebbah (Ed.), *Sciences de la société*, 101.
- Fang, Y.-J. (1994). "Riots" and Demonstrations in the Chinese Press: A Case Study of Language and Ideology. *Discourse & Society*, 5(4), 463-481.
- Gabrielatos, C. & Baker, P. (2008). Fleeing, Sneaking, Flooding. A corpus Analysis of Discursive Constructions of Refugees and Asylum Seekers in the UK Press, 1996-2005. *Journal of English Linguistics*, 36(1), 5-38.
- Holmes, S. & Castañeda, H. (2016). Representing the "European refugee crisis" in Germany and beyond: Deservingness and difference, life and death. *American Ethnologist*, 43(1), 12-24.
- Kaufmann, L. (2006). Les voies de la déférence. Sur la nature des concepts sociopolitiques. *Langage et société*, 117, 89-115.
- Krieg-Planque, A. (2009). *La notion de "formule" en analyse du discours*. Besançon: Presses Universitaires de Franche-Comté.
- Kosnick, K. (2014). Mediating Migration: New Roles for (Mass) Media. *InMedia*, 5, 271-334.
- Kyriakides, C. (2016). Words don't come easy. Al Jazeera's migrant-refugee distinction and the culture of (mis)trust. *Current Sociology*, 65(7), 933-952.
- Palmer, M. (2006). Nommer les nouvelles du monde, in J. Arquembourg, G. Lochard and A. Mercier (coord.). *Hermès*, 46, 47-56.
- Searle, J. (1995). *The Construction of Social Reality*. New York: Simon & Schuster.
- Siblot, P. (2001). De la dénomination à la nomination. Les dynamiques de la signifiante nominale et le propre du nom. *Cahiers de praxématique*, 36, 189-214.
- Sinclair, J. (1991). *Corpus Concordance Collocation*. Oxford: Oxford University Press.
- Stubbs, M. (1996). *Text and Corpus Analysis*. Oxford: Blackwell.
- Valluy, J. (2005). La nouvelle Europe politique des camps d'exilés: genèse d'une source élitaine de phobie et de répression des étrangers. *Cultures & Conflits*, 57, 13-69.
- Van Dijk, T. (1991). *Racism and the Press*. London/New York: Routledge.
- Van Dijk, T. (2006). Discourse and manipulation. *Discourse & Society*, 17(3), 359-383.
- Veniard, M. (2013). Du profil lexico-discursif de crise à la construction du sens social d'un événement. In Londei, D., Moirand, S., Reboul-Touré, S. & Reggiani, S. (Eds.). *Dire l'événement. Langage, mémoire, société* (pp. 221-232). Paris: Presses Sorbonne Nouvelle.