Danielle Leeman,Université Paris Ouest Nanterre La Défense & UMR 5191 ICAR (CNRS & ENS/Lyon 2)

Achille Falaise,UMR 5191 ICAR (CNRS & ENS/Lyon 2)

Les prépositions devant les noms de région et de département français

# Thématique, problématique, hypothèse

L'usage des prépositions devant les noms géographiques (désignant des pays, des régions, des départements, des villes...) est généralement résumé ainsi (nous limitons l'extrait à ce qui concerne les noms de région et de département, qui constituent l'objet de la présente contribution) :

« Les noms de provinces se construisent avec *en* lorsqu'ils sont féminins ou qu'ils commencent par une voyelle (*dans* est possible). Quand ils sont masculins et à initiale consonantique, ils prennent *en* ou *dans le* ; rarement *au*[[1]](#footnote-1). /.../ Devant les noms de département français, on emploie *dans* etl'article[[2]](#footnote-2) /.../ Certains noms acceptent aussi *en* sans article : les noms composés singulier (*En Haute-Marne,* *En Seine-et-Marne*) ; certains noms simples, sans doute confondus avec les noms des anciennes provinces (*En Dordogne, En Gironde, En Vendée, En Vaucluse, en Aveyron*). *En* l'emporte avec *Savoie* (ancienne province) et *Corse* (comme nom d'île) » (Grevisse & Goosse 2008 : § 1051).

Deux problèmes se posent à la lecture de telles descriptions : d'une part, on ne voit pas le rapport (d'ailleurs non explicité) entre le choix de la préposition, le genre du nom qu'elle introduit ou la voyelle qui le commence[[3]](#footnote-3) – ces dernières propriétés n'étant, de fait, jamais invoquées lors de la définition des prépositions. D'autre part, la, disons, « prudence » des formulations « *en* ou *dans le* », « rarement », « certains noms » ne permet pas d'être sûr du choix que l'on opère dès que l'on sort des exemples cités : peut-on dire, en parlant de la région Centre, aussi bien *En Centre* que *Dans le Centre* ? Les noms composés de département *Puy-de-Dôme* ou *Tarn-et-Garonne* permettent-ils *En Puy-de-Dôme* ou *En Tarn-et-Garonne*[[4]](#footnote-4) ? Les noms *Gers*, *Finistère* ou *Calvados* ne correspondent pas à celui d'une ancienne province, que penser alors d'attestations telles que[[5]](#footnote-5) :

*Dans le vieil Auch, logé dans l'ancien couvent des jacobins, le musée conserve une très belle collection de pièces précolombiennes, légués par des passionnés du pays. Au sous sol des vestiges de l'Auch antique et médiévale. A l'étage, des pièces retracent la vie* ***en Gers*** *au XIXème et début du XXème siècles.  
A voir découvrir cette belle région, la vie et les centres d'intérêt de ses habitants.* [[6]](#footnote-6)

[Finistère Rando | Randonnées en Finistère](https://encrypted.google.com/url?q=http://www.finistere-rando.com/&sa=U&ved=0ahUKEwiF_-j7kbbNAhXKXBoKHWDtCb4QFghYMA4&usg=AFQjCNFCB_REFapGFI9ZOD5tpOVvPG5fXw)

www.**finistere**-rando.com/‎

Un guide nature, des paysages magnifiques, une faune et une flore variées... c' est la recette gagnante des balades nature proposées en **Finistère**.

Note : 4 - Critique de voyageur TripAdvisor - 30 mai 2013

Kashmir: Enfin un restaurant indien convenable **en Calvados**; - consultez 79 avis ... les meilleures offres et comparez les prix pour Caen, **France** sur **TripAdvisor**.

Selon une étude conduite par Céline Vaguer, que nous remercions pour son aide constante et amicale, sur le journal *Le Monde* des années 1987, 1990, 2000, 2010, la formulation *en Bretagne* est représentée 497 fois mais *dans la Bretagne* 3 fois (ce qui corrobore la parenthèse de Grevisse et Goosse concernant les noms féminins *supra*) ; en revanche, *en Limousin* apparaît 49 fois contre *dans le Limousin* 8 fois, *en* *Rhône-Alpes* 121 fois et *dans* *le Rhône-Alpes* 31 fois – ce qui montre que la probabilité d'occurrence après l'une ou d'autre préposition n'est pas équivalente.

Néanmoins, la grammaire citée, en mentionnant par deux fois « les anciennes provinces » pour expliquer l'emploi inattendu de *en*,met sur une voie d'explication[[7]](#footnote-7) que nous adopterons comme hypothèse : cette préposition active dans le nom locatif qu'elle introduit un stéréotype, en ceci qu'il ne s'agit pas tant d'évoquer un lieu (comme avec *à* ou *dans*[[8]](#footnote-8)) qu'un trait pertinent  (ou plusieurs) aux yeux d'une certaine collectivité ; en l'occurrence, donc, *en* serait employé lorsqu'il s'agit d'un lieu dont on convoque l'identité qui lui est propre : celle de la province à laquelle il a appartenu historiquement.

## Provinces, départements, régions : brève situation historique

Sans entrer dans le détail de l'Histoire, l'évolution de l'organisation administrative du territoire français peut se résumer selon trois grandes époques (Joly, 2016) : d'abord, la Gaule est partagée en « provinces », caractérisées par leurs particularités géographiques, historiques, culturelles (costume, gastronomie, musique, danse, langue...), leurs peuples ; les provinces deviennent les fiefs des seigneurs sous l'Ancien Régime[[9]](#footnote-9) ; ensuite, la Révolution substitue aux provinces les « départements » – entités administratives dont les dénominations sont fondées sur des choix géographiques (rivière comme *Doubs*, *Drôme*, montagne comme *Puy de Dôme, Jura*...), sans doute pour éviter la reprise des identités féodales ; enfin dans les années cinquante est entreprise et continuée, au fil des années soixante-dix, la création des « régions », fondées sur les départements et comme eux instaurations administratives, pour lesquelles la plupart des Conseils régionaux adoptent les anciens noms de province (ce qui conforte l'hypothèse retenue). La loi sur la décentralisation (1982) accroît leurs prérogatives et l'évolution se poursuit sous la Présidence de François Hollande avec la réduction de leur nombre.

## Le corpus constitué pour tester et valider l'hypothèse

Afin d'étudier la distribution des prépositions en fonction des noms de région et de départements français, nous avons constitués trois corpus à partir du Web :

1. Un corpus diachronique issu du journal *Le Monde*, formé à partir des premiers paragraphes de 10 % des articles, diffusés gratuitement par le journal sur le Web, pour la période 1950-2015 (16,5 M mots).
2. Un corpus contemporain de commentaires touristiques, extraits du site *TripAdvisor* (4,5 M mots).
3. Un corpus contemporain issus de sites Web d'offices de tourisme (722 offices, 3,6 M mots).

De ces trois corpus, seul *Le Monde* est diachronique, et nous l'utiliserons donc principalement. Le corpus *Tourisme* servira à comparer les résultats obtenus pour *Le Monde* sur la période 2010-2015 avec le discours touristique, populaire (*Commentaires*) et institutionnel (*Offices de tourisme*).

Ces pages Web ont été téléchargées, après extraction des textes et des métadonnées pertinentes pour notre étude, puis annotées morphosyntaxiquement et lemmatiquement ; les corpus ainsi obtenus ont été intégrés à nos outils d'étude de corpus (ScienQuest [ref] et TXM [ref]).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *Le Monde* | | | | | | | *Tourisme* | |
| 1950-1959 | 1960-1969 | 1970-1979 | 1980-1989 | 1990-1999 | 2000-2009 | 2010-2015 | *Commentaires* | *Offices de tourisme* |
| 0,9 M | 1,4 M | 1,6 M | 1,6 M | 1,8 M | 4,7 M | 4,2 M | 4,5 M | 3,6 M |

Nombre de mots par corpus.

Ont ensuite été prélevées de ces corpus les séquences {*à*, *dans*, *en*} + nom de région ou de département français, les occurrences de ces noms pouvant s'écrire avec et sans traits d'union (*Rhône-Alpes* et *Rhône Alpes*), et la préposition *à* étant susceptible d'apparaître sous les formes amalgamées *au* et *aux*. Ces résultats sont filtrés manuellement ; sont exclus tous les cas où la préposition est contrainte par le verbe : « *Laisser tout le mérite à la Bretagne*»*,* «*le transfert de pouvoirs législatifs à la Corse*», ou bien où la préposition fait partie d'un nom propre : « *Saint-Quentin-en-Yvelines* », « *Rock en Seine* ».

# Préposition *en* et noms de régions

La liste des régions (en 2014)[[10]](#footnote-10) comporte des noms féminins et masculins – parmi ces derniers aucun ne commence par une voyelle. La dénomination *Mayotte* n'est pas précédée d'article[[11]](#footnote-11).

## Les noms de région féminins

Le corpus *Le Monde* contient 655 occurrences de la séquence {*à*, *dans*, *en*} + nom de région féminin. La description donnée pour les provinces par *Le Bon Usage* se vérifie sur ce corpus : les noms féminins se trouvent quasi exclusivement avec *en*. Selon ces attestations, la « règle » est donc bien *en* + N*.* Cela se vérifie aussi sur le corpus *Tourisme*, qui ne comporte que 3 occurrences (sur 346) de *en* + nom de région féminin, dans des cas très particuliers, comme (au sujet de la cathédrale de Strasbourg) : « *Point de repère* ***dans l'Alsace*** *(on la voit depuis l'aéroport), elle est à la fois imposante et pleine de finesse.* », où l'Alsace est plutôt présentée sur le mode topographique que régional.

Figure 1: Fréquence relative (%) des prépositions en fonction de la décennie pour les noms de régions féminins, sauf Guadeloupe, Martinique, Mayotte, Réunion (546 occurrences)

.Par contre, le nom des îles de *La Réunion* et *Mayotte* se construit toujours avec *à* (34 occurrences). Le premier cas peut s'expliquer par le fait que *en* n'est pas compatible avec l'article – or le déterminant *la* fait partie de la dénomination[[12]](#footnote-12). Cependant, *Mayotte* s'emploie sans article et n'est pas pour autant compatible avec *en –* il y a donc une autre explication à trouver.

## Guadeloupe et Martinique

Le cas de la Guadeloupe et de la Martinique est remarquable. Ces deux régions, qui sont aussi des îles, ont changé de préposition au cours du XXe siècle (Reboul, 1994). Ainsi, *Guadeloupe* est attesté principalement (> 85%) avec la préposition *en* depuis les années 1980, mais celle-ci faisait jeu égal avec *à* dans les années 1960 et 1970. Encore plus net : *Martinique*, nom toujours introduit par la préposition *à* dans les années 1950, ne se construit plus qu'avec *en* depuis les années 1990.

Figure 2: Fréquence relative (%) des prépositions en fonction de la décennie pour Guadeloupe (41 occurrences − dont une seule pour 1950)

Figure 3: Fréquence relative (%) des prépositions en fonction de la décennie pour Martinique (34

occurrences).

|  |  |
| --- | --- |
|  |  |
| Le corpus Offices de tourismes ne contient pas d'occurrence de Guadeloupe ou Martinique, elles sont par contre nombreuses dans les données contemporaines de Commentaires touristiques : 305 en Martinique (92%) pour 25 à la Martinique (8%), rejoignant ainsi la tendance observée sur Le Monde – 59 % de ces commentaires émanent de personnes se déclarant originaires de la Martinique, mais les 25 à la Martinique émanent toutes d'européens (Français métropolitains ou Belges), donc pour qui l'île n'est pas constitutive de leur identité, ce qui étaye notre hypothèse. | |

## Conclusion sur les noms de région féminins

En conclusion, si l'on peut expliquer que, quoique féminin, le nom *La Réunion* ne soit pas précédé de *en*, du fait de l'incompatibilité entre la préposition et la présence obligatoire de l'article, un problème subsiste : en admettant la pertinence de la justification qui vient d'être avancée, le nom *Mayotte* devrait pouvoir être introduit par *en*, ce qui n'est pas le cas – il ne l'est que par *à.*

En revanche, l'évolution de *à la Guadeloupe* et *à la* *Martinique* vers la généralisation de *en Guadeloupe* et *en Martinique* peut s'expliquer selon notre hypothèse initiale, à savoir comme l'effet d'un changement de perception et de représentation de ces îles, qui recouvrent et revendiquent désormais une identité spécifique assumée et reconnue comme telle – perspective ouverte par la politique de décentralisation qui restitue aux régions une autonomie, et donc le choix de leur image, par conséquent la possibilité de rendre visible une culture propre – ; dans cette optique, le fait que l'appellation *Guyane* est toujours précédée de *en* signifierait que ce territoire est durablement doté d'une figure particulière, stéréotypique : par exemple celle d'un lieu où sont envoyés les forçats dans des bagnes aux conditions de vie inhumaines (ce depuis la fin du XVIIIe siècle, avec la Révolution française).

## Les noms de région masculins

Le cas de la région Centre a été laissé de côté car, dans le corpus retenu, la forme, si elle est extrêmement fréquente, correspond rarement au nom de la région – elle n'en est pas moins attestée. D'autre part, contrairement aux noms de région féminins, les noms de région masculins sont, sauf *Limousin*, des noms composés, artificiels, qui ne peuvent pas être confondus avec des noms de territoires anciens comme *Alsace* ou *Bretagne*. Il n'est donc pas étonnant de rencontrer une seule occurrence de nom de région masculin, précisément avec ce nom (*dans le Limousin*), avant l'établissement des 22 régions « classiques » en 1970. L'étude *infra* porte donc uniquement sur la période 1970-2015. Elle montre que, globalement, ces noms s'emploient avec *dans* ou avec *en* (85 apparitions en tout), mais, observation qui confirme les dires du *Bon Usage*, aucune suite avec *à* n'est attestée, en dehors des trois verbes qui imposent cette préposition (*consacrer à*, *donner à*, *s'arrêter à*), que nous avons donc écartés. Constat qui étaye notre hypothèse : le graphe montre une forte progression de la préposition *en* entre la décennie 1970 et la période 2010-2015 (figure ci-dessus), passée de 13 % à 90 % des occurrences.

Figure 4: Fréquence relative (%) des prépositions en fonction de la décennie pour les noms de régions masculins : Languedoc-Roussillon, Limousin, Midi-Pyrénées, Nord-Pas-de-Calais, Pays de la Loire, Poitou-Charentes, Rhône-Alpes entre 1970 et 2010 (85 occurrences, pas d'occurrence avant 1970).

Les noms de région *Midi-Pyrénées* (10 occurrences), *Poitou-Charentes* (5 occurrences) et *Rhône-Alpes* (18 occurrences) ne se rencontrent jamais, dans notre corpus, avec la préposition *dans*, mais toujours avec la préposition *en*, et ce même dans les années 1970. Une hypothèse d'explication possible est que ces substantifs se distinguent par le fait que le deuxième nom du mot composé est féminin (cette préposition étant liée à ce genre pour les noms de pays, comme on sait). Certes, pour *Pays de la Loire*, nos données n'illustrent pas d'emploi avec *en* mais cette construction est attestée, par exemple *PME en Pays de la Loire* (sur un site Web de la CCI des Pays de la Loire[[13]](#footnote-13)), ou encore *L'Europe s'engage en Pays de la Loire*, présent sur de nombreux sites Web financés par l'Union Européenne.

Sauf *Limousin* (du moins au sein de notre corpus), les noms de région masculins (*Languedoc-Roussillon*, *Midi-Pyrénées*, *Nord-Pas-de-Calais*, *Pays de la Loire*, *Poitou-Charentes* et *Rhône-Alpes*), passent tous progressivement d'une introduction par *dans* (unique en 1970) à des constructions avec *en* (entre 70 et100% en 2010). Le petit nombre d'occurrences pour chaque région prise individuellement ne permet pas de faire des comparaisons précises entre elles, mais, dans tous les cas, la progression de *en* au détriment de *dans* est nette et indubitable : on passe d'une construction privilégiée avec *dans* à hauteur de 87 % dans les années 1970, à une distribution qui favorise *en* (82 %) dans les années 2010, comme le montre la figure 4 ci-dessous pour les deux noms de régions les plus fréquents : *Languedoc-Roussillon* (11 occurrences) et *Nord-Pas-de-Calais* (20 occurrences).

Figure 5: Fréquence relative de en par rapport à dans pour deux noms de région, par période de 20 ans entre 1970 et 2010 (pas d'occurrences 

avant 1970).Le cas de *Limousin* est plus difficile à traiter dans le cadre de nos données, qui n'en procurent que 16 occurrences, et aucune pour la décennie 2010 : sur la période 1970-2009, ce nom de région se répartit à part égales entre *dans* (50%)et *en* (50%), sans évolution : il se peut que ce fonctionnement différent provienne du fait que *Limousin* constitue un cas particulier parmi les noms de région masculins, étant le seul nom identique à celui d'une ancienne province.

## Conclusion sur les noms de région masculins

La description du *Bon Usage* s'avère ici à modifier : sauf pour *Limousin* (mais les résultats diffèrent selon les corpus pris en considération : les chiffres et pourcentages auxquels a abouti Céline Vaguer (citée *supra*) ne recouvrent pas ceux qui ressortent de nos données), *en* et *dans* ne se partagent pas de manière égale l'emploi devant les noms de région masculins puisque, soit *en* est la préposition unique attestée, soit une évolution passe d'une domination de *dans* à une large majorité de *en*. Cette progression étaye l'hypothèse retenue en ceci que l'évolution peut s'interpréter comme un effet de la décentralisation entamée dans les années cinquante et concrétisée par la loi du 2 mars 1982, permettant aux régions d'assumer leur autonomie et, par là, de revendiquer leur identité propre.

Si l'on compare le comportement de *à*, *en*, *dans* devant les noms de pays et devant les noms de région (en l'occurrence les régions françaises, telles que politiquement et administrativement définies), on ne peut manquer d'être surpris par le fait que le fonctionnement soit différent : la préposition *à* est de rigueur devant les noms de pays masculins (commençant par une consonne), mais elle n'est pas attestée devant les noms de région masculins. La préposition *en* dévolue aux noms de pays féminins et aux noms de pays masculins commençant par une voyelle se généralise désormais, au détriment de *dans*, devant tous les noms de région – fussent-ils masculins. La préposition *dans* ne s'accommode pas des noms de pays ni des noms de région féminins non spécifiés, mais elle se trouve devant les noms de région masculins non spécifiés (\* *Dans le Portugal, il fait beau* vs *Dans* *le Limousin / Vaucluse, il fait beau*).

# Prépositions et noms de département

Les noms féminins seront traités d'abord, les noms masculins dans un deuxième temps. Pour les deux séries, en suivant les informations fournies par *Le Bon Usage*, il sera d'abord vérifié si les noms peuvent être introduits par *dans*, ensuite si les noms composés sont les seuls à pouvoir être précédés de *en*, et enfin si les noms simples attestés avec *en* peuvent aussi être les appellations d' « anciennes provinces ». Le corpus est représenté dans les deux listes suivantes ; la périodisation retenue couvrant les années cinquante, y ont été intégrés d'anciens noms de départements (*Basses-Alpes*, *Basses-Pyrénées*, *Côtes-du-Nord*, *Loire-Inférieure*, *Seine*, *Seine-Inférieure*).

**ATTENTION : il manque les deux listes !!!**

Les noms *Guadeloupe*, *Guyane*, *Martinique*, *Mayotte* et *Réunion*, qui sont aussi des noms de région et ont déjà été traités précédemment, sont ici écartés, de même que *Paris*[[14]](#footnote-14), tant il se confond avec la ville de Paris – autrement dit, ne sont prises en compte que les formes qui correspondent avec certitude à un nom de département : ne sont ainsi pas non plus retenues des occurrences comme « *Dans le Nord et l'Est*», « *À Tripoli, dans le Nord*»(le Nord de la Libye, et non le département français du Nord), ou encore *« coulé à pic dans l'Hérault*».

## Les noms de département féminins

À travers les 788 occurrences relevées sur notre corpus, apparaissent trois cas de figure, selon que les noms sont toujours introduits par *dans*, ou toujours construits avec *en*, ou montrent une évolution diachronique allant de *dans* à *en*.

**Sont toujours attestés avec *dans*:**

Ardennes, Aube, Aude, Basses-Alpes, Basses-Pyrénées, Bouches-du-Rhône, Deux-Sèvres, Drôme, Eure, Hautes-Alpes, Landes, Marne, Nièvre, Oise, Orne, Pyrénées-Atlantiques, Pyrénées-Orientales, Sarthe, Seine-Inférieure, Somme, Vienne, Vosges, Yonne, Yvelines.

**Apparaissent toujours avec *en* :**

Charente, Corrèze, Corse-du-Sud, Haute-Marne, Indre-et-Loire, Loire-Inférieure, Lozère, Mayenne, Saône-et-Loire, Savoie, Seine-Maritime.

**Apparaissent avec dans et en :** voir figure 6.

Rappelons que, selon le *Bon Usage* (*op. cit*.), « devant les noms de département, on emploie *dans* et l'article » : le corpus montre que ce n'est en rien systématique puisque neuf d'entre eux ne se construisent qu'avec *en* – ceque l'ouvrage n'avait signalé que pour *Savoie* – et quinze avec les deux prépositions, sans qu'il s'agisse forcément de noms composés (*Meurthe, Moselle*...) ni de noms d'anciennes provinces[[15]](#footnote-15) (*Charente, Corrèze, Creuse*...). Grevisse & Goosse justifient la distribution *en Corse* par le fait que *Corse* est un nom d'île, mais c'est aussi le nom d'une ancienne province – c'est cette explication qui sera retenue ici, car elle traite cette formulation de la même façon qu'elle interprète toutes les autres[[16]](#footnote-16).

Figure 6: Noms de départements féminins qui apparaissent avec en et dans

.Figure 7: Fréquence relative (%) des prépositions en fonction de la décennie pour les noms de département féminins (788 occurrences).

Globalement, on observe aussi une progression, dans le temps, de l'emploi de *dans* vers l'emploi de *en*, moins marquée que pour les régions mais tout de même marquant une inversion des proportions en 1980 – *dans* baisse progressivement au profit de *en* qui finit par devenir majoritaire, le cas de *Isère* est à cet égard particulièrement clair. Notre hypothèse est que, comme pour les noms de région, cette évolution, à partir des années soixante-dix, de l'emploi de la préposition devant les substantifs désignant les départements correspond à une évolution des mentalités liée à la politique de décentralisation entamée dans la seconde moitié du XXe siècle, laquelle entraîne pour chaque région et pour chaque département une revendication identitaire, le désir d'avoir une image spécifique se substituant au pur découpage administratif qui, à l'inverse, gomme toute particularité.

Figure 8: Fréquence relative (%) des prépositions en fonction de la décennie pour Isère (31coccurrences)

## Les noms de département masculins

Figure 9: Évolution de la fréquence des noms de département 

masculins (384 occurrences).Notre corpus n'atteste aucune occurrence avec *en* avant 1970. Ensuite, s'observe une courbe caractéristique, avec l'introduction de *en* en 1970, suivi d'une forte croissance en 1980, puis une stabilisation à un niveau plus bas à partir des années 1990. La courbe ci-dessous montre que ceux-ci se construisent toutefois toujours très majoritairement avec *dans*. De fait, au sein de nos données, seuls quelques noms se construisent avec *en* : *Aveyron* (8 occurrences), *Finistère* (21 occurrences), *Loir-et-Cher* (5 occurrences), *Lot-et-Garonne* (9 occurrences) et *Tarn-et-Garonne* (4 occurrences)[[17]](#footnote-17).

Sur le corpus *Tourisme*, seules 6 occurrences de noms de département masculins se construisent avec *en*. Le plus fréquentest *Finistère* (15 occurrences, dont 2 avec *en* et 13 avec *dans*), ainsi que quelques autres, absents des données issues du journal *Le Monde*, notamment *Maine-et-Loire* (2 occurrences dans *Tourisme*, toujours avec *en*). Mais, comme vu précédemment, le Web témoigne de nombreuses occurrences de *en Calvados*, attesté dans aucun de nos corpus, par exemple, un article de *Ouest-France* intitulé *Pompiers payants en Calvados*[[18]](#footnote-18). L'emploi de cette préposition pour *Calvados* est d'ailleurs attesté plusieurs fois dans ce journal, quoique de manière très marginale : le moteur de recherche interne du site du journal retourne ainsi 19 pages de résultats pour *en Calvados*, contre 690 pour *dans le Calvados*.

Figure 10: Noms de départements masculins qui apparaissent avec en et dans

.Figure 11: Évolution de la fréquence des noms de département masculins Aveyron, Finistère, Loir-et-Cher, Lot-et-Garonne, Tarn-et-Garonne (47 occurrences)

## Conclusion générale sur les données observées

Se référer à de grands corpus systématiques permet de saisir des phénomènes linguistiques et des évolutions discursives que les théories ou les descriptions partielles (à partir d'un certain nombre de données jugées représentatives), *a fortiori* les exemples forgés, ne peuvent prévoir (Blumenthal 2006, 2008 ; Vigier 2016) – mais tout corpus présente des limites : ainsi avons-nous pu observer que des attestations absentes de notre corpus principal (*Le Monde*), se trouvaient en nombre dans d'autres genres textuels (corpus *Tourisme*), ou attestés dans d'autres documents disponibles sur le Web. En l'état de la recherche toutefois, il nous paraît raisonnable de considérer validée l'hypothèse adoptée, selon laquelle, d'une part, il existe une évolution dans l'emploi des prépositions *en* et *dans* devant les noms de région et de département, et ce au profit de *en* qui, progressivement, occupe le terrain auparavant détenu par *dans*; et selon laquelle, d'autre part, cette mutation discursive est l'écho d'un changement dans les représentations sociétales – en l'occurrence lié au remaniement géographique de l'organisation du territoire français décidé par les politiques à partir des années cinquante, qui autorise les nouvelles entités dites « régions » à renouer avec leur identité d'antan dans la mesure où elles correspondent au « anciennes provinces » – mouvement qui se répercute sur les départements constituant ces régions.

Pour soutenir le bien-fondé de cette hypothèse ratifiée empiriquement, il reste néanmoins à expliquer pourquoi le processus de substitution s'opère dans un sens (de *dans* à *en*) et non dans l'autre (de *en* à *dans*), ce qui suppose la définition de l'identité sémantique des prépositions concernées (section 4 ci-dessous). Cette explicitation permet de résoudre le double problème précédemment resté en suspens : (1) la préposition *à* est de rigueur devant les noms de pays masculins commençant par une consonne, elle n'est en revanche pas attestée devant les noms de région de même nature (*La vie est belle au Brésil / Portugal / Maroc* vs \* *au Limousin / Vaucluse / Nord-Pas-de-Calais*), et (2) la préposition *dans* ne s'accommode pas des noms de pays ni des noms de région féminins non spécifiés, mais se trouve devant les noms de région masculins non spécifiés (\* *Dans le Portugal, il fait beau* vs *Dans* *le Limousin / Vaucluse, il fait beau*).

# L'identité des prépositions et son impact en discours

Eskénazi (*op. cit.*) rappelle que, dans le cadre théorique guillaumien postulant le mode de genèse du mot en langue (c'est-à-dire avant sa production effective en discours)[[19]](#footnote-19), deux étapes se succèdent : la première, à partir d'une idée vague de ce que l'on cherche à exprimer – il s'agit d'une saisie « précoce », qui relève de « l'avant » –, est de progressivement préciser, singulariser, l'objet en question ; la deuxième, à l'inverse, est de l'ordre de la généralisation (le mot précédemment identifié ne sert pas seulement à exprimer la notion spécifique qui en est la source mais s'en abstrait afin de pouvoir énoncer toutes les notions du même type, les conceptualiser) – la saisie est alors de l'ordre de « l'après ».

Dans cette organisation générale du système linguistique, une relation existe entre, d'un côté, la préposition *en*, le genre féminin et la voyelle – qui relèvent d'une saisie d'avant, première, relativement floue, intuitive et subjective, de la notion à verbaliser –, et, d'un autre côté, la préposition *à*, le genre masculin et la consonne – qui relèvent d'une saisie d'après, donc élaborée, abstraite, conceptuellement construite. Cette hypothèse permet de rendre compte du fait que les noms de pays féminins, ou masculins commençant par une voyelle, se construisent avec *en*, et que les noms de pays masculins commençant par une consonne se construisent avec *à* (Leeman, 2015b).

Cependant le corpus de la présente étude ne suit pas cette organisation, puisque de plus en plus de noms masculins de région ou de département sont introduits par *en*, qu'ils commencent par une voyelle ou une consonne, et qu'aucun n'est construit avec *à*. Il est donc nécessaire d'approfondir la définition de l'identité des prépositions pour être en mesure d'avancer une hypothèse résolvant cette contradiction.

## L'identité de la préposition *à*

Sémantiquement, la préposition *à* fait du nom de lieu une simple localisation, assez imprécise dans la mesure où elle est susceptible de renvoyer à un lieu spécifique aussi bien que non spécifique – ainsi, *Elle part à* *l'université* peut se comprendre comme « elle est mutée dans l'enseignement supérieur » ou comme « elle va sur son lieu de travail » –, impliquer ou non une activité – par exemple « *Être à la fenêtre*, pour un être humain, c'est sans doute être debout « près » d'une fenêtre, mais aussi et surtout être en train de « regarder » par la fenêtre » (Katz, 2002 : 37-38) tandis que *Être à la maison* n'engage rien de tel –, et n'emporte pas de précision sur le lieu en ceci que *Être au café*, par exemple, ne dit pas exactement où la personne se trouve : au bar ? Dans la salle ? Aux toilettes ? En terrasse ? Voire devant l'établissement ?

L'exclusion de la préposition *à* par les noms de région et de département peut alors s'expliquer par le fait que, dans la représentation que l'on a de ces territoires, ces derniers ne peuvent être bornés à une simple localisation vague, car ils sont étroitement liés à la vie concrète des individus – interprétation qui va dans le même sens que celle du développement de la préposition *en*. Cette hypothèse peut s'étendre à la résolution du problème précédemment resté en suspens, qui est que, si la préposition *à* est de rigueur devant les noms de pays masculins commençant par une consonne, elle n'est en revanche pas attestée devant les noms de région de même nature (*La vie est belle au Brésil / Portugal / Maroc* vs \* *au Limousin / Vaucluse / Nord-Pas-de-Calais*). Ce qui signifierait, selon l'ensemble des hypothèses adoptées jusqu'ici sur le sens de la préposition *à*, que le « pays » (au sens de « nation ») constitue un lieu relativement flou, lointain, abstrait[[20]](#footnote-20) (par opposition à la région ou au département, qui concernent de plus près la vie personnelle des individus), parce qu'il s'agit d'une entité plus vaste, dont la définition peut résulter d'un rassemblement forcé de peuples différents et dont l'identité doit être l'objet de connaissances acquises (en particulier à l'école).

Selon notre hypothèse issue de la théorie guillaumienne résumée *supra*, le fait que cette construction avec la préposition *à*, donccette représentation linguistique que nous postulons, ne concerne que les noms de pays masculins commençant par une consonne est due au fait que, comme *à*, le masculin et la consonne relèvent de « l'après ».

**4. 2 L'identité de la préposition *en***

La préposition *en* n'a pas pour première vocation d'exprimer la localisation, même prise en un sens dit « spatial » : comme l'indique Guillaume (1919), elle extrait du nom qu'elle introduit une qualité qui « se reverse » sur l'identité de celui qui la précède ; ainsi l'annonce suivante produit-elle un double effet interprétatif :

[Job d'été : le travail en usine (témoignage) - digiSchool](http://www.google.fr/url?q=https://www.digischool.fr/job-ete/job-dete-le-travail-en-usine-4529.php&sa=U&ved=0ahUKEwjs85Dap7bNAhUDnRoKHUy9AHYQFghaMAk&usg=AFQjCNF-np0cErhSTbwoREM7lKoN-HPFoA)

https://www.digischool.fr/job.../job-dete-le-travail-**en-usine**-4529.php‎

26 avr. 2011 **...** L'un des travails les plus demandés est celui du travail **en usine**. Justine l'a testé durant deux années et nous fait part de son expérience.

D'une part, la relation avec l'usine introduite par la préposition spécifie rétroactivement l'identité d'abord affectée à l'entité « travail » et *a fortiori* « job d'été », mais d'autre part aussi permet réciproquement de renouveler l'idée que l'on se fait habituellement d'une usine. La suite *en N* assure donc un double rôle prédicatif (Leeman & Vaguer 2014), requalifiant aussi bien ce qui la précède que ce qui la suit (Franckel, 2015 ; Paillard, 2015). La question est alors de déterminer en quoi consiste cette qualification.

La préposition *en* relève, on l'a dit, d'une saisie « précoce » : à cette étape, il s'agit pour la pensée (et la langue par laquelle elle fonctionne) de définir, de singulariser un objet d'abord appréhendé de manière floue, intuitive, subjective, en fonction de connaissances et ressentis personnels. Ainsi le nom géographique introduit par *en* représente-t-il non une entité objective, abstraite, rationnellement élaborée[[21]](#footnote-21), mais, au contraire, une réalité affective appréhendée à travers un vécu concret : la perception et la mémoire que l'on a de son relief, son climat, son histoire, ses coutumes, sa cuisine, sa musique et ses chants, son patrimoine, sa langue (ou son accent en parlant le français), ses gens dans leur spécificité – conçue de manière impressionniste (et non scientifique) [[22]](#footnote-22)... – en bref, ces patries locales que constituaient les « anciennes provinces » dont parlent Grevisse & Goosse.

Dans cette même perspective, l'emploi des noms de région (et de département) masculins avec la préposition *en* s'interprète comme l'écho, dans les perceptions des territoires, de la réforme de décentralisation entreprise dans les années cinquante, qui, conférant aux régions une autonomie et des prérogatives nouvelles, leur permet du même coup de se prévaloir d'une identité spécifique, de même que les départements qui les constituent prétendent désormais au statut de « terroir »[[23]](#footnote-23).

**4. 3 Résolution du problème posé par *à Mayotte* vs \**en Mayotte***

L'hypothèse a été avancée plus haut que *en* ne se combine pas avec *La Réunion* du fait de la présence de l'article, qui fait partie de l'appellation[[24]](#footnote-24) – à l'opposé des noms tels que (*la*) *France* : si l'on dit bien *un parfum qui vient de France*, ce n'est pas le cas de \**un parfum qui vient de Réunion*[[25]](#footnote-25). Mais il reste à expliquer pourquoi *Mayotte* non plus n'admet pas *en*, alors que le nom n'est pas précédé d'un déterminant. Notre supposition est que, si les noms de région féminins non spécifiés ne se combinent pas avec la préposition *dans* (comme les noms de pays) et si *dans La Réunion* n'apparaît pas non plus attesté, c'est que, linguistiquement, ces endroits sont susceptibles d'être perçus comme des pays[[26]](#footnote-26) – et non, seulement, comme des régions ou des départements français[[27]](#footnote-27) (sur les représentations liées aux noms de pays, voir Cislaru 2006). En revanche, la construction *dans Mayotte* se rencontre (certes beaucoup moins fréquemment que *à Mayotte*), comme en témoignent :

*Et nous avons confiance, car partout dans Mayotte, au coin de chaque village*[[28]](#footnote-28) …

*ces nouveaux clandestins condamnés au déracinement pour des raisons économiques et qui, lorsqu'ils parviennent à s'infiltrer dans Mayotte*...[[29]](#footnote-29)

*Ici, les naturalistes (des scientifiques qui font des relevés sur tout, dans Mayotte) estiment qu'il reste 5% de la foret primaire.*[[30]](#footnote-30)

ce dont on peut conclure que, contrairement à La Réunion et aux autres régions ou départements, Mayotte n'est pas conçue ou concevable (dans le système linguistique français) comme un pays – mais seulement comme une simple localisation. Confirme cette analyse le fait que le nom se combine aussi avec *à* mais non avec *en*, donc ne se prête pas à l'extraction d'une identité susceptible de caractériser une autre entité.

**4. 4 Résolution des problèmes posés par *dans***

La préposition *dans* ne s'accommode pas des noms de pays (masculins ou féminins) ni de certains noms de région féminins s'ils ne sont pas spécifiés : \* *Dans le Portugal, on aime les pâtisseries* vs *Dans* *le Portugal de mon enfance, on aimait les pâtisseries,* à opposer à *Dans le Limousin / Vaucluse, on aime les pâtisseries*. Du côté des noms féminins, apparaissent non acceptables aussi bien \**Dans la France / \*Dans l'Alsace, on aime la* *bière*[[31]](#footnote-31)(*Dans l'Ariège* est naturel (sans détermination) mais c'est un nom de département et non de région). La théorie guillaumienne permet d'apporter des réponses partielles à ces questions:

1. selon cette conception, on l'a vu *supra*, *dans* relève de « l'après », donc est du côté du masculin – il est par conséquent normal qu'il introduise un nom masculin (non spécifié) – ; le fait que ce ne soit pas le cas des noms de pays (même masculins) signifierait que ces derniers n'évoquent pas dans le système linguistique et dans les représentations des locuteurs une réalité précise à laquelle on se sent attaché (contrairement au nom de région), au sein de laquelle on trouverait une localisation subjectivement pertinente (voir note 20) – analyse cohérente avec celle que l'on a menée pour la préposition *à* ;
2. le nom féminin relève de l'avant, donc est *a priori* compatible avec *en*, mais non avec *dans* ; cependant, s'il est spécifié, donc singularisé, de ce fait conçu plus précisément et par conséquent saisi tardivement, il devient compatible avec *dans*, qui relève de l'après (*Dans l'Alsace de ma jeunesse*).

En conclusion : de par leur place dans l'organisation du système linguistique, *dans* (préposition « d'après ») ne s'accommode pas des noms de pays ni des noms de région féminins non spécifiés (qui relèvent de « l'avant ») ; si elle peut néanmoins se trouver devant les noms de région masculins non spécifiés (\**Dans le Portugal, il fait beau* vs *Dans* *le Limousin / Vaucluse, il fait beau*), c'est que ces substantifs sont en eux-mêmes spécifiques, au sens où ils ont une identité précise (combinant géographie, histoire, culture, etc.) ; ils sont donc compatibles avec la préposition *dans* qui, dans le parcours allant de la saisie première, d'avant, floue, à la saisie dernière, d'après, qui singularise l'objet, se situe dans cette phase ultime, après *à* [[32]](#footnote-32).

# 5. Conclusion

La complexité de la description des noms géographiques est bien connue (Maurel *et al.*, 2014), et a donné lieu à de nombreuses descriptions, en particulier pour définir leur introduction par une préposition, mais sans qu'une explication soit fournie pour expliquer les distributions observées. L'objectif est ici de déterminer le comportement des noms de région et de département français, après avoir étudié celui des noms de pays (Leeman 2015). L'analyse d'un corpus diachronique de 21 millions de mots constitué à partir du journal *Le Monde*, de commentaires publiés sur le site de *TripAdvisor*, et de sites Web d'offices de tourisme, complété par des attestations trouvées sur le Web [à l'aide de Google], permet de détecter une évolution marquée par le développement de *en* au détriment de *dans*, correspondant à la réorganisation du territoire français depuis les années cinquante, ce qui montre que l'emploi des prépositions reflète les représentations de la société. Le recours à la théorie guillaumienne permet d'expliquer les possibilités et impossibilités combinatoires, ainsi que les mutations dans les emplois.

Mots clés : préposition, noms géographiques, corpus, théorie guillaumienne

**Références**

Blumenthal, P. (2006) « De la logique des mots à l'analyse de la synonymie », *Langue française* 150 : 14-31.

Blumenthal, P. (2008) « Combinatoire des prépositions : approche quantitative », *Langue française* 157 : 37-51.

Bottineau, D. (2010) « Typologie de la déflexivité », *Langages* 178 : 89-114.

Cislaru, G. (2006) « Noms de pays et autoreprésentation dans le discours des périodiques nationaux français, anglophones, roumanophones et russes », *Les Carnets du CEDISCOR* 9 : 131-144.

Constant, M. (2010) « Prépositions locatives et noms propres géographiques », *Cahiers du Cental* 6 : 73-80.

Eskenazi, A. (1987) « Député DE Saône-et-Loire – Préfet DU Rhône – En Vendée », *LINX* 16 : 28-69.

Flament-Boistrancourt, D. (1993) « *Café de Colombie* ou *Café de la Colombie ?* », in *Le FLE de 0* *à 300 heures*, Angers, CIDEF : 91-112.

Franckel, J.-J. (2015) « La préposition *en* dans les mots du discours. Le cas de *en plus* », *Scolia* 29 : 79-96.

Garrigues, M. (1993) « Prépositions et noms de pays et d'îles : une grammaire locale pour l'analyse automatique des textes », *Lingvisticae Investigationes* XVII-2 : 281-306.

Grevisse, M. & A. Goosse (200814) *Le Bon Usage*, Bruxelles, De Boeck & Duculot.

Joly, A. (2016) *L'Imposture occitane à la sauce béarnaise*, Toulon, Editions des Dauphins, collection hors-série de *Modèles* *linguistiques* « Langues et cultures patrimoniales de France ».

Katz, E. (2002) « Systématique de la triade spatiale *à, en, dans* », *Travaux de Linguistique* 44 : 35-50.

Leeman, D. (2015a) « La préposition *en* et les noms de pays », dans *Phraséologie et profils combinatoires : lexique, syntaxe et sémantique – Hommages offerts à Peter Blumenthal*, Paris, Champion : 189-200.

Leeman, D. (2015b) « Hypothèse de résolution du problème posé par l'emploi des prépositions devant les noms de pays », Colloque *SynSémLex*, Sorbonne, mars, actes en 2016 chez Benjamins.

Leeman, D. & C. Vaguer (2014) « La préposition peut-elle être prédicative ? Le cas de la préposition *en*», *Verbum* XXXVI-2 : 397-420.

Maurel D., Baron M., Bouchou-Markhoff B., Vitas D. (2014). Prolexbase. A Multilingual Relational Database of Proper Names. *Cahiers de Linguistique* 40(2):49-71. hal-01119318

Paillard, D. (2015) « Les locutions *en + N* dans leurs emplois comme marqueurs discursifs », *Scolia* 29 : 97-114.

Reboul, S. (1994) « *A la Guadeloupe / en Guadeloupe* : vers une interprétation cognitive ? », *Langue française* 103 : 68-79.

Vigier, D. (2013) « Sémantique de la préposition *en* : quelques repères », *Langue française* 178 : 3-20.

Vigier, D. (2016) « Dialogue autour de « *en*» et «*dans*» : apports d’une perspective diachronique », Journées *L'argument, la circonstance et la particule. Quelqu'un a parlé de système ?*, Paris, Sorbonne, 21-22 juin, actes à par.

1. Sont donnés entre autres exemples : *En Normandie, en Anjou, Dans l'Anjou, Dans l'Artois, En Périgord, en Dauphiné, En Limousin, en Berry, en Poitou, Dans le Limousin, Dans le Berry.* [↑](#footnote-ref-1)
2. Exemples fournis : *Dans la Gironde, Dans l'Ain, Dans la Seine-Maritime, Dans le Bas-Rhin, Dans le Vaucluse, Dans les Bouches-du-Rhône.* [↑](#footnote-ref-2)
3. C'est évidemment aussi le cas des noms de pays, traité dans Leeman 2015a et 2015b. [↑](#footnote-ref-3)
4. On peut lire sur des panneaux autoroutiers « Vous entrez en Puy-de-Dôme » ou, sous le panneau d'interdiction concernant les véhicules de plus de 3t5 : « Transit interdit en Tarn-et-Garonne ». On pourrait invoquer l'économie de caractères que représente *en* par rapport à *dans le*, mais il existe aussi « Transit interdit dans le département de la Vienne », qui montre que le petit nombre de lettres n'est pas le souci premier des rédacteurs de ces panneaux. [↑](#footnote-ref-4)
5. Sans compter les messages générés automatiquement. [↑](#footnote-ref-5)
6. Commentaire déposé en juin 2015 sur le site de Trip Advisor. [↑](#footnote-ref-6)
7. Peut-être à la suite du remarquable travail d'A. Eskenazi (1987), dont les grandes lignes sont présentées dans Leeman *op. cit.* [↑](#footnote-ref-7)
8. Les trois prépositions sont définies et distinguées dans Leeman *op. cit*. [↑](#footnote-ref-8)
9. La liste est fournie sur le site : http://www.francogene.com/rech-fr/prov-fr.php [↑](#footnote-ref-9)
10. <http://www.cartesfrance.fr/carte-france-region/carte-france-regions.html#ixzz446atH8gG> [↑](#footnote-ref-10)
11. Le chapitre 1er de la Loi organique no 2007-223 du 21 février 2007 considère *Mayotte* comme un nom féminin, au vu de l'Article LO6111-3 : « Mayotte est représentée au Parlement et au Conseil économique et social dans les conditions définies par les lois organiques. ». [↑](#footnote-ref-11)
12. Selon la Commission nationale de toponymie, l'article est intégré au nom, ce qui explique la dénomination *Région*  *de La Réunion* par opposition à *Région Guadeloupe, Région Guyane, Région Martinique*. Les justifications de l'appellation *La Réunion* présentent le point commun de la faire émaner du nom commun. [↑](#footnote-ref-12)
13. [http://www.pmepaysdelaloire.fr](http://www.pmepaysdelaloire.fr/) [↑](#footnote-ref-13)
14. Le nom est masculin, donc normalement incompatible avec *en* (Leeman, 2015a). [↑](#footnote-ref-14)
15. Voir la liste des anciennes provinces fournie en Annexe. [↑](#footnote-ref-15)
16. De surcroît, les noms d'île n'ont en rien un comportement homogène : le fait de désigner une île ne paraît pas pertinent pour l'emploi du nom, même si *île* peut fonctionner comme « classifieur » : *l'île de La Réunion* (Garrigues, *op. cit*., Constant, 2010). [↑](#footnote-ref-16)
17. Le recours à d'autres sources montre cependant d'autres attestations, comme vu *supra* avec *Gers*. [↑](#footnote-ref-17)
18. <http://www.ouest-france.fr/normandie/caen-14000/pompiers-payants-en-calvados-choquant-et-lamentable-reagit-la-cgt-4125293> [↑](#footnote-ref-18)
19. « La psychomécanique du langage partage avec le générativisme au moins un objectif : rendre compte des conditions de production de l'énoncé (performance / actualisation du discours) par une architecture mentale dynamique constitutive d'un savoir-faire cognitif (compétence / langue comme systèmes de représentation) » (Bottineau, 2010 : 91). [↑](#footnote-ref-19)
20. Les « pays » sont, de fait, la plupart du temps le résultat de victoires militaires ou d'accords politiques établis selon des critères qui ne tiennent pas vraiment compte de l'identité des populations concernées : en témoigne particulièrement, au XXe siècle, le découpage imposé par la fin des colonisations lors de l'obtention de l'indépendance par les anciennes colonies. [↑](#footnote-ref-20)
21. C'est pourquoi, de leur création jusqu'au milieu du XXe siècle, les noms de département ne sont pas introduits par *en* (car ils relèvent alors d'une création administrative rationnelle, non fondée sur le patrimoine historique), mais par *dans*, qui présente le lieu comme une entité circonscrite l'intérieur de laquelle sont localisées d'autres entités. [↑](#footnote-ref-21)
22. Le Breton est têtu, l'Auvergnat près de ses sous, le Marseillais vantard, le Corse paresseux... [↑](#footnote-ref-22)
23. C'est ainsi que l'on peut interpréter le développement de *en Gers* et autres signalés *supra*. [↑](#footnote-ref-23)
24. Contrairement à ce qu'indique M. Constant (2010 : 73). [↑](#footnote-ref-24)
25. Voir à ce propos Flament-Boistrancourt (1993). [↑](#footnote-ref-25)
26. Le mot *pays* est d'ailleurs utilisé pour les désigner : *un vin de pays*, *le pays bigouden*, *avoir le mal du pays.*.. [↑](#footnote-ref-26)
27. Il en va de même pour la Guadeloupe, la Guyane et la Martinique, dont les noms (sans spécification) sont rarement ou ne sont pas combinés avec *dans*. [↑](#footnote-ref-27)
28. <https://www.cairn.info/revue-enfances-et-psy-2007-2-page-147.htm> [↑](#footnote-ref-28)
29. <http://resf37.free.fr/spip.php?article291> [↑](#footnote-ref-29)
30. <http://www.routard.com/forum_message/1030736/8/quelques_tuyaux_pour_ceux_qui_veulent_venir_a_mayotte.htm> [↑](#footnote-ref-30)
31. On relève néanmoins des attestations de noms de région féminins introduits par *dans,* par exemple en l'occurrence sur Trip Advisor : « Un séjour inoubliable dans l'Alsace grâce à mes amis des 3 pierres », « Parfait pour famille en visite dans l'Alsace ». [↑](#footnote-ref-31)
32. Comme vu précédemment, *être au café*  ne précise pas l'endroit exact où l'on se trouve, tandis que *être dans le café* constitue une localisation intérieure. De plus, *il va au café* peut s'interpréter aussi bien de façon générique que d'une façon spécifique, tandis que *il va dans le café* reçoit une signification uniquement spécifique. Ainsi, dans le parcours qui va de l'avant à l'après, *en* se situe avant *à,* qui se trouve avant *dans*. [↑](#footnote-ref-32)