



HAL
open science

Le Lexique comme reflet des représentations sociales : ses variations selon le corpus considéré

Danielle Leeman, Max Silberztein

► **To cite this version:**

Danielle Leeman, Max Silberztein. Le Lexique comme reflet des représentations sociales : ses variations selon le corpus considéré. *Linx*, 2020, 80. halshs-02885062

HAL Id: halshs-02885062

<https://shs.hal.science/halshs-02885062v1>

Submitted on 30 Jun 2020

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

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

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

Max Silberztein, Besançon & ELLIADD

Le lexique comme reflet des représentations sociales : ses variations selon le corpus considéré.

Résumé.

L'analyse distributionnelle des mots relevant du champ de l'habitation, comme *appartement*, *maison*, *studio*, montre qu'ils sont munis, en discours, de connotations différentes : ils reflètent la manière, parfois surprenante, dont la société se représente leur référent, mais cette représentation est susceptible de varier selon le discours. La valeur d'un lexème ne peut donc être saisie qu'au terme de l'analyse de corpus de genres divers.

Mots clés.

Lexique - analyse distributionnelle – représentation sociale – corpus - genre

Abstract.

The distributional analysis of French words that belong to the *habitation* [habitat] field, such as *appartement*, *maison*, *studio*, shows that they have different connotations when used in a given discourse, as they reflect the sometime surprising manner of how society sees the referents of these concepts. These representations vary according to the discourse in question; the value of a lexeme can thus be defined only by studying its occurrences in corpora of various genres.

Keywords.

Lexicon – distributional analysis – social representation – corpus - genre

Ce qui domine sans doute, dans la mémoire collective, concernant l'apport de Jean Dubois et sa situation dans la linguistique française, c'est, en matière de syntaxe et de sémantique, la priorité donnée à l'observation des formes, des propriétés distributionnelles et transformationnelles, dans la ligne tracée par le structuralisme harrissien des années cinquante. Or si dès sa thèse, éditée en 1962 – et par conséquent dont la préparation a dû commencer un lustre auparavant –, Jean Dubois revendique effectivement l'organisation du Lexique en système, il précise que ces rapports sont « historiquement formés au cours de l'évolution de la langue et communs à la collectivité » – précision intéressante en tant qu'elle s'écarte de la doxa du structuralisme *stricto sensu*, lequel voit dans la langue une synchronie à dissocier de son histoire : pour Jean Dubois, en revanche, le Lexique ne saurait être hypostasié de la société ni des personnes qui la constituent car l'identité de ses unités ne provient pas seulement du système d'oppositions qu'est la langue mais aussi de l'usage des mots

par ceux qui la parlent – usage pénétré de la vision qu'ils ont du monde et des événements qu'ils vivent.

La présente contribution attire l'attention sur une observation qui rejoint ce qui était alors au fondement de la thèse de Jean Dubois mais qui a également fondé sa théorie de l'analyse du discours – laquelle ne devait à Harris que sa méthodologie, et non ses objets (les discours inscrits dans la réalité sociale¹) ni ses objectifs (l'hypothèse est que, selon leur source et leur destinataire, les discours s'opposent par leur structure et le sens attribué à leur lexique²).

Les distributions sont révélatrices de l'identité sémantique des unités lexicales, et la divergence dans la probabilité des co-occurrences permet la détermination de sous-classes (Z.S. Harris 1954 trad. 1970), ouvrant la possibilité d'affiner la description syntaxique et d'établir une classification sémantique (Dubois 1979), prémisses de la dite « théorie des classes d'objets » qui émerge au début des années quatre-vingt-dix (G. Gross 1994) et méthode appliquée à des corpus exhaustifs, systématisée dans *Les Verbes français* entre autres (Dubois et Dubois-Charlier 1997). Ce sur quoi nous voulons attirer ici l'attention, c'est que ces distributions, dans un champ lexical donné – par exemple celui de l'habitation –, ne sont pas exploitées de la même manière dans les discours effectifs, reflétant la représentation sociale dont est l'objet le référent du nom et témoignant donc de l'identité sémantique de l'unité lexicale correspondante en langue. Si l'on compare ainsi les occurrences de *en appartement*, *en maison* et *en studio* telles qu'attestées sur les premières pages de *GOOGLE* à la suite de ces requêtes « *en + N* » (Leeman 2017), *en appartement* se révèle apparaître généralement dans un contexte négatif dans la mesure où il s'agit de résoudre une difficulté :

*Paul se retrouve en appartement avec sa copine*³

quel chien en appartement ?

Comment vivre avec son chat en appartement ?

*Les galeries d'art s'installent en appartement*⁴

Poêle à bois en appartement

Le lombricompostage en appartement

Les inconvénients de la vie en appartement

1 Par exemple les discours pédagogiques (produits par l'institution, les enseignants, les élèves), les discours pathologiques, les discours politiques.

2 Ainsi, dans le domaine politique, les discours de Blum (socialiste) et de Thorez (communiste), v. Courdresses (1971).

3 Le verbe *se retrouver* implique le résultat d'une situation problématique.

4 Parce qu'en louer ou en acheter une revient beaucoup trop cher.

Vivre seul en appartement

trois enfants et vivre en appartement

Vivre en appartement, quelles dépenses ?

tandis que *en studio* (lorsqu'il s'agit du lieu d'habitation) exploite aussi un trait négatif mais en général celui de l'exiguïté :

Ces trucs auxquels j'ai renoncé depuis que je vis en studio

Vivre seul en studio

Habiter en studio impose de maîtriser l'art de l'organisation de l'espace.

et *en maison* n'intervenant que dans des noms composés désignant des lieux fonctionnels de type *maison de retraite, maison d'hôte, maison de repos, maison relais* ou *maison de ville* – ou n'intervenant que par contraste avec *en appartement* dans des énoncés du type *Vivre en maison ou en appartement ?*.

Néanmoins, si les attestations relevées dans *GOOGLE* peuvent révéler des tendances intéressantes, ce corpus ne permet pas de généraliser la recherche en procédant par catégorie grammaticale, par exemple selon le modèle *dans Dét (Adj) appartement*. C'est pourquoi il nous a paru nécessaire d'utiliser un outil linguistique d'exploration de corpus. Il ressort de cette expérience la conclusion que les représentations qui se dégagent de tel type de corpus (en l'occurrence les attestations extraites de *GOOGLE* citées *supra*) ne convergent pas nécessairement avec celles qui ressortent de tel autre (ici le journal *Le Monde* de l'année 2016) : les représentations sociales sont donc sujettes à variation selon les genres textuels.

CONSTITUTION DU CORPUS ET MÉTHODE D'OBSERVATION

Pour des raisons de place, on s'est limité au sens locatif des noms choisis, précédés de trois prépositions seulement ; le corpus a été constitué en extrayant du corpus du journal *Le Monde* pour l'année 2016 (21 411 829 mots) :

- la préposition *à* suivie d'un déterminant, suivi éventuellement d'adjectifs antéposés, suivis d'un des cinq noms cités (est repéré par exemple *à la maison*) ;
- la préposition *dans* suivie d'un déterminant, suivi éventuellement d'adjectifs antéposés, suivis d'un des cinq noms cités (entre autres *dans ce petit studio*) ;
- la préposition *en* suivie éventuellement d'adjectifs antéposés, suivis d'un des cinq noms cités, comme dans *en beaux appartements*.

Sont ensuite éliminées les occurrences non pertinentes pour l'étude : celles où les noms ne désignent pas un espace où l'on est susceptible d'habiter, comme lorsque *studio* désigne un lieu de tournage ou d'enregistrements musicaux.

L'observation consiste à comparer, d'une part, le nombre d'occurrences des noms et leur pourcentage afin de déterminer ce qui apparaît le plus familier dans la vie des locuteurs (du moins à travers l'image qu'en donne le discours du journal *Le Monde*), ainsi que, d'autre part, les adjectifs ou compléments accompagnant éventuellement le nom ; ces derniers sont essentiellement de trois types : « dimensionnels » (en tant qu'ils qualifient les dimensions du référent), « axiologiques » (en tant qu'ils portent un jugement sur le sentiment que le référent produit sur le locuteur), « objectif » (en tant qu'ils dénotent une propriété observable). De même, prendre en compte le cotexte est susceptible de renseigner sur les représentations sociales des référents de ces noms, par l'indication des activités qui leur sont associées.

1. Comparaisons quantitatives

Sur un total de 965 occurrences pertinentes pour notre analyse, *maison* en représente 512 soit 53%, *appartement* 341 soit 35%, *villa* 59 soit 6%, *chambre* 221 dont, pour ce qui nous intéresse, *chambre d'hôtel* 30 soit 3%, *chambre d'hôpital* 6 soit 0,6% et *studio*, *chambre de bonne* ou *chambre de service* 5 soit 0,5% – *chambre d'hôte* n'apparaissant qu'une fois.

La forte proportion du nom *maison* n'est pas *a priori* surprenante, ce nom étant susceptible de fonctionner comme hyperonyme ; on n'est pas non plus étonné de voir *appartement* bien davantage cité que *villa*, en songeant que beaucoup plus de gens logent dans des immeubles ; en revanche, il paraît inattendu de voir *studio* si peu représenté⁵, dans la mesure où l'on aurait tendance à penser que bon nombre de jeunes, en particulier, vivent dans ce type de logement. De fait, selon un tableau statistique de 2013⁶ inventoriant le nombre de gens occupant des « un ou deux » pièces (nous n'avons pas trouvé de chiffres concernant ceux qui vivent en studio) :

- sur les 22,4% des ménages propriétaires occupants de leur logement dans un immeuble collectif, 27% le sont d'un ou de deux pièces – autrement dit près de 6% du total des propriétaires occupants ;
- sur les 75% des ménages locataires de leur logement dans un immeuble collectif, 44,9% le sont d'un ou de deux pièces – autrement dit plus de 29% ;

Dans les deux cas, on est loin du 0,5% de présence du nom *studio* dans le corpus considéré : si, dans ce dernier, la représentativité de *maison* et *appartement* reflète l'importance de leur référent dans la réalité sociale (selon les statistiques consultées), il n'en va pas de même pour *studio* et autres logements d'une pièce qui n'y reçoivent, littéralement, qu'un traitement quantitativement négligeable.

2. Occurrences de *chambre (de bonne, de service, d'hôtel)*

5 Le corpus ne contient qu'une occurrence de *studette*, une occurrence de *TI*, et aucune de *FI*.

6 <https://www.logisneuf.com/statistique-immobiliere.html#parc-logement>

Les seuls termes qui nous intéressent (en tant qu'ils désignent un endroit où on loge) sont *chambre de bonne*, *chambre de service*, *chambre d'hôpital* et *chambre d'hôtel*. 38 de leurs occurrences (soit 93%) sont construites avec la préposition *dans* et 3 occurrences avec la préposition *en* (*en chambre d'hôtel*) ; seul l'adjectif *petit* est représenté et uniquement dans la séquence *une petite chambre de bonne*.

La requête « chambre » sur *GOOGLE* (le 14 juin 2019) aboutit d'abord à une publicité pour des chambres d'hôtes, puis, massivement, à des réclames concernant le mobilier ou la décoration de chambres à coucher, et finalement, la mention de possibilités de réservation de chambres d'hôtel. Comme vu plus haut, le corpus du *Monde*, à l'inverse, comporte d'abord davantage d'occurrences de *chambre d'hôtel*, *chambre d'hôte* n'arrivant qu'en dernier.

Sur *GOOGLE*, il faut spécifier « chambre d'hôpital » pour obtenir des informations centrées sur cette thématique, et dans ce cas, les informations qui apparaissent concernent la question de savoir où acquérir et à quel prix une chambre d'hôpital (ses différentes composantes), les insuffisances actuelles des chambres d'hôpital, les évolutions en cours et les améliorations à en attendre pour les patients et les soignants – le discours s'adresse donc plutôt à l'administration gérant l'institution. Dans le corpus du *Monde* (année 2016), le contexte est tout autre, s'agissant de circonstances dramatiques concernant les patients :

Ilyes, trois ans, était installé dans la chambre 17. Il était arrivé avec ses parents dans l'après-midi pour une angine.

Ce n'est qu'après sa mort que, dans la chambre 17, on découvre que le flacon qui a servi à la perfusion est du chlorure de magnésium, en lieu et place du B46, du chlorure de sodium.

Nous retrouver ne sera pas simple, confie-t-il dans sa chambre d'hôpital. On était cinq et on n'est plus que deux.

qu'il entendit sur un poste à galène, dans sa chambre d'hôpital, avant de s'éteindre le surlendemain des suites d'une crise cardiaque.

Le gouvernement indien a invoqué cette loi pour confiner Irom Sharmila dans une chambre d'hôpital d'Imphal, la capitale du Manipur, où elle était alimentée de force par sonde nasale plusieurs fois par jour.

Les prédicats concernant *chambre de bonne* diffèrent également selon le type de corpus : la requête « chambre de bonne » débouche dans *GOOGLE* sur la définition du vocable, sur des annonces de vente, de location, des conseils (ne pas investir dans l'achat d'une chambre de bonne, comment aménager un studio ou une chambre de bonne) et des témoignages (avantages ou inconvénients de vivre dans une chambre de bonne, prix locatifs abusifs, propriétaires bailleurs sans scrupules). Dans le corpus *Le Monde*, la chambre de bonne est petite (une occurrence) ou minuscule (deux occurrences) et le plus généralement occupée gratuitement, éventuellement en échange

de ménage ou autres services – il est ici question de vie concrète dans des conditions précaires :

une chambre de bonne sans chauffage et sans loyer, mise à disposition par un vieux bourgeois veuf

meurt à 88 ans, en 2004, ruiné, dans une chambre que les nouveaux propriétaires du Copacabana Palace lui offriront gracieusement

s'installe dans une chambre de bonne prêtée contre de petits travaux par la comtesse de Blégiers

La requête « chambre de service » aboutit dans *GOOGLE* et dans le corpus du *Monde* à un résultat à peu près semblable à ceux de « chambre de bonne » (définition, annonces de vente ou de location, conseils ou appels à l'aide, témoignages dans un cas, et dans l'autre : *Logée chez sa patronne dans une chambre de service*) – dans les textes (comme précédemment), *chambre de service* peut être relayé aussi bien par *studio*, *studette* ou *chambre de bonne*.

3. Occurrences de *studio*

Toujours introduit pas la préposition *dans*, éventuellement qualifié de *petit*, le *studio* dans le corpus du *Monde* est très généralement lié à des qualificatifs négatifs (mais qui ne sont pas spécialement liés à l'exiguïté du lieu, comme c'est le cas dans les attestations relevées dans *GOOGLE*, voir *supra*) : situé dans un immeuble *vétuste*, *sentant la clope et la litière de chat*, occupé par une *prostituée* (2 occurrences), *spartiate* ou *sans charme*, situé au *rez-de-chaussée* (position peu prisée, par opposition aux étages élevés), il peut être la scène d'événements déplaisants :

L'homme qui n'accepte pas que l'on débarque dans son studio pour pisser sur son tapis...

ou dramatiques :

... jetant désespérément un enfant en contrebass. Les cris ont fini par résonner toute la nuit dans le studio du couple,

Le *studio* peut aussi être attribué gratuitement à des réfugiés :

La veille, Yosef a pu emménager dans un petit studio payé par la ville.

4. Occurrences de *villa*

Ce terme apparaît 59 fois dans le corpus (mais jamais après la préposition *en*, contrairement à ce que l'on observe dans *GOOGLE*) :

— 51 occurrences après la préposition *dans* ;

— 8 occurrences après la préposition *à*, liées à des verbes de déplacement (*mener, escorter, se présenter, accompagner, retourner*).

Les adjectifs épithètes sont surtout axiologiques : *cossue* (2), *huppée, fleurie, tranquille, belle* (2), *luxueuse, ancienne, jolie, somptueuse, splendide, superbe, vaste* (ce dernier adjectif relève aussi du dimensionnel).

Abandonnée, blanche, décorée, effondrée, rénovée, vieille sont objectifs.

Grande (6), *immense* sont dimensionnels, de même que *vaste* (qui est aussi axiologique).

► Les qualificatifs convergent pour conférer à *villa* l'évocation d'un haut degré de fortune et / ou d'une classe sociale supérieure. Les prépositions sont cependant purement locatives, *à* et *dans* ne faisant pas du nom le support d'un statut comme le ferait *en*. Sur *GOOGLE* en revanche, *en villa* est attesté dans les discours publicitaires mettant en valeur un certain type de location touristique, le plus souvent de luxe, qui s'oppose à d'autres catégories de séjours (*appartement en villa, location en villa, etc.*).

La villa, dans le corpus du *Monde*, apparaît le plus généralement dans un contexte de luxe spécifié par la présence d'une piscine, l'association à des palais, la situation géographique, la vue, la tranquillité :

dans des villas dominant les collines toscanes
dans des villas tranquilles d'autres quartiers de Marseille
dans la belle villa avec piscine prêtée par des artistes
dans leur villa de bord de mer à Lacanau
dans sa superbe villa sur les hauteurs de Cassis

► C'est donc un lieu caractérisé par des habitants de haut niveau social, qui coulent des jours heureux entre fêtes et concerts :

dans l'ancienne villa de villégiature de la Marquise de Montebello
dans les villas des dignitaires de l'EI
organise une fête pour son soixantième anniversaire dans sa villa de Goa, et offre à ses invités un concert privé

l'entrepreneur était surnommé « le roi du bon temps », à cause des fêtes légendaires organisées dans ses villas

ou bien un lieu où l'on peut vivre tranquille, à l'abri :

La comtesse se réfugie dans sa vaste villa pour se mettre à l'abri des bombardements

Il vivait reclus dans une immense villa qui surplombait Hollywood

► Mais, de par ses caractères précédemment énumérés, la villa est aussi l'objet de cambriolages ou de crimes et donc d'enquêtes :

Il s'agit de juger les accusés pour la séquestration violente, dans une villa d'Aix-en-Provence, du frère d'un buraliste

poignardée à mort avec quatre autres personnes dans sa villa de Los Angeles.

un cambriolage simulé dans sa grande villa pour accéder au coffre.

Le feu est à peine éteint dans la villa où l'EI détenait ses prisonniers

En conclusion, le point commun aux corpus *GOOGLE* et *Le Monde* est la connotation luxueuse du nom *villa*, qu'officialise la préposition *en* dans *GOOGLE* dans les discours adressés aux voyageurs et vacanciers, lesquels acquièrent un statut particulier par un séjour *en villa*. Les événements liés à la villa sont plus sombres, voire tragiques, dans le corpus du *Monde* : s'il peut s'agir du lieu de fêtes liées à la fortune de son propriétaire, la villa est le plus souvent la cible de cambriolages, d'incendies, le lieu de trafics, de séquestrations ou de meurtres.

5. Occurrences de *appartement*

Les 341 séquences qui contiennent le terme *appartement* se répartissent en :

- 318 occurrences avec la préposition *dans* (93 %) ;
- 15 occurrences avec la préposition *à*, dont 10 sont imposées par la construction du verbe (*conduire, livrer, passer, avoir accès, mener* (2), *correspondre, attacher, donner droit*) ;
- 8 occurrences avec la préposition *en*, dont 4 sont le fait de la construction du verbe *transformer* – donc imposées.

Autant dire que le nom est librement construit essentiellement avec la préposition *dans*.

40 séquences contiennent des adjectifs antéposés, dont *petit* est le plus fréquent :

- 31 occurrences d'un des adjectifs dimensionnels suivants (78%) : *grand, petit, immense, spacieux* – *spacieux* est à la fois dimensionnel et axiologique ; *petit*, avec ses 24 occurrences, est de loin le plus utilisé (77,4% des adjectifs dimensionnels) ;
- 9 occurrences d'un des adjectifs axiologiques suivants : *bel, humble, joli, lumineux, luxueux, pauvre, spacieux, somptueux* – *spacieux* est à la fois dimensionnel et axiologique.

Les adjectifs postposés peuvent être les mêmes, mais laissent apparaître l'existence d'appartements sans confort (nous les soulignons) : *cossu, design et cosy, rongé par la moisissure, douillet, niché dans une petite rue tranquille, lumineux, tout simple, à loyer modéré, bourgeois, confortable, haussmannien* (2), *luxueux, minuscule, sans eau chaude, vétuste, grand style, ancien.*

► Construit avec *en*, le nom *appartement* apparaît associé à l'enfermement ou à une vie marginale, la seule occurrence positive *beaux appartements* apparaissant en contraste avec *squats glauques* et d'ailleurs étant imposée par la construction de ... *en* :

/les victimes accumulent/ compulsivement et jusqu'à la saleté des objets en appartement – une autre forme d'enfermement

/relation trouble entre les deux hommes/. Huis clos à deux personnages en appartement ils parlent peu, évoluent la nuit, de squats glauques en beaux appartements parisiens

► La thématique de l'enfermement se retrouve assez fréquemment dans les combinaisons avec *dans* où l'appartement est susceptible d'indiquer le refuge de fuyards, la cachette de délinquants, d'objets volés ou d'activités illicites :

les deux hommes que les services belges surveillent, en ce début janvier, dans ce petit appartement

Dans cet appartement, les enquêteurs ont trouvé une Kalachnikov,

Partout dans la ville se cachent des bordels très discrets, situés dans de petits appartements.

Les produits étaient stockés au fils des arrivages dans des appartements des cités

Ils sont soupçonnés d'avoir hébergé le fugitif dans l'appartement de la rue des Quatre-Vents

► Mais l'enfermement n'est pas nécessairement destiné à enfreindre la loi :

Je ne voulais pas devenir paranoïaque, vivre cloîtrée dans l'appartement

Les habitants, réfugiés chez eux, lumières éteintes dans les appartements, sont apparus aux fenêtres

la tragédie des retraitées grecques confinées dans leurs appartements

Depuis trois mois, il vivait reclus dans son appartement de la via Della Panetteria

Calfeutrée dans son appartement de Sao Paulo, au Brésil,

L'avocat avait pendant l'Occupation défendu et caché dans son appartement des membres de la Résistance danoise.

Maria Callas, retirée des scènes et cloîtrée dans son appartement parisien

En résumé, l'appartement apparaît connoté négativement dans le corpus *Le Monde* comme dans les attestations relevées dans *GOOGLE*, mais la connotation négative repose sur une base différente dans les deux cas : plutôt l'exiguïté dans *GOOGLE*, plutôt l'enfermement et le fait de (se) cacher dans le corpus du *Monde*.

6. Occurrences du nom *maison*

Le corpus contient 1.501 séquences construites avec le mot *maison*.

Ont été exclus de la concordance les noms propres, par ex. *Maison Blanche*, *Maison de la culture*, *maison Lenôtre*, *maison de Molière*, y compris lorsque ces noms propres apparaissent en minuscules (ex. *maison blanche*), ainsi que de nombreux noms d'institutions ou d'entreprises : *maison d'arrêt*, *maison centrale*, *maison close*, *maison de correction*, *maison de courtage*, *maison de couture*, *maison de disques*, *maison d'édition*, *maison de torture*, *maison mère*, *maison théâtre*, etc. Ont aussi été exclues les occurrences qui contiennent les prépositions *à* et *en* lorsqu'elles sont appelées par la construction du verbe (*donner à*, *transformer en*), telles :

... *on pourrait donner une autre dimension à cette maison bicentenaire...*

... *transformer leurs baraques en bois en petites maisons de brique...*

Nous avons cependant gardé les termes composés qui désignent des endroits où l'on peut habiter : *maison de campagne*, *maison d'enfance*, *maison de famille*, *maison d'hôte*, *maison de maître*, *maison de plain-pied*, *maison de repos*, *maison de retraite*, *maison de santé*, *maison conjugale*, *maison éclusière*, *maison familiale*, *maison individuelle*, *maison mitoyenne*, *maison natale*, *maison paternelle*.

Les 512 occurrences formant le corpus analysé comportent :

— 115 séquences *à la maison*, plus une occurrence de la préposition *jusqu'à* précédant le démonstratif :

...*se rendre jusqu'à cette grosse maison posée sur la place de Crisser...*

Il n'y a pas d'autre emploi de la préposition *à* locative (toutes les autres séquences prépositionnelles en *à* correspondent à des compléments d'objet indirect en *à*).

— 20 constructions avec la préposition *en*, toutes figées : 17 constructions contiennent un terme composé d'institution : *en maison d'hôte*, *en maison de repos*, *en maison de retraite*, *en maison de santé*. L'analyse ne retient pas les 2 occurrences de la construction semi-figée : *de maison en maison*, et *de maison vide en maison vide*.

— 395 (77 %) constructions avec la préposition *dans*.

La plupart des modificateurs sont des compléments de nom locatifs, comme dans :

... *et vit avec Alison, dans une maison au bord de l'Hudson...*

... *nombreuses œuvres sont installées dans sa maison du Luberon, aménagée par son compagnon...*

mais non tous :

...plantation de palmiers dattiers, puis à la maison d'un cousin...
... ouvriers qui, chez eux, vivent dans de petites maisons de terre ou de bois...

Les adjectifs modificateurs antéposés ou postposés représentés sont :

- des adjectifs de lieu ou de nationalité, tels *florentine, limougeaude, palestinienne* ;
- des adjectifs dimensionnels, comme *grande, grosse, immense, petite* ;
- des adjectifs axiologiques : *agréable, ancienne, belle, cossue, déglinguée, enchantée, fétiche, futuriste, jolie, moderne, modeste, mystérieuse, normale, superbe, triste, vieille* ;
- des adjectifs « objectifs », par exemple : *abandonnée, blanche, décrépite, délabrée, flottante, grise, humide, imaginaire, isolée, jaune, jumelle, nouvelle, vide* ;
- des adjectifs supposant une appréciation « sociale » : *bourgeoise, célèbre, privée*.

6.1. **Interprétation des distributions de en maison**

Sur les 17 occurrences recensées, nous négligerons la seule occurrence de *maison d'hôte* et les trois occurrences de *maison de santé*, vocables qui nous paraissent trop peu représentatifs, au profit des 12 occurrences de *maison de retraite* – en remarquant toutefois que cette thématique nous semble vraiment sous-représentée dans ce corpus, au regard de l'importance sociale qu'elle peut présenter dans le monde contemporain. Ces noms composés sont les compléments de mots (verbes, participes passés, noms) indiquant l'installation (*départ, hébergés, placement, séjours*) dont le sujet, s'il est explicite, concerne le patient (6 occurrences) : *le réalisateur, je, les personnes âgées, les personnes, mes parents, de nombreux seniors*, plutôt que le personnel (3 occurrences) : *filles (discute de son travail), employée précaire, professionnels*.

La préposition *en* (conformément à l'identité que lui attribue l'hypothèse guillaumienne, Guillaume 1919) active donc, parmi les traits du nom, celui qui définit un statut à celui qui l'habite : ici, le patient qui y est placé et, très secondairement, à celui qui y travaille.

6.2. **Interprétation des distributions de à la maison**

Comme précédemment, ne sont pas pris en compte les cas où la préposition fait partie de la construction du verbe, comme dans *relier N à N* ou *conférer N à N*.

L'association *à la maison* forme un bloc dont le sens équivaut à « chez soi » où le pronom s'accorde à la situation énonciative (*Je rentre à la maison = chez moi / Nous arrivons à la maison = chez nous*) et vaut quel que soit le type d'habitation, l'inverse n'étant pas vrai : *Viens à la maison* peut se dire même si le locuteur habite un appartement, *Viens à l'appartement* ne peut pas se dire si le locuteur vit dans une maison.

- Le syntagme *à la maison* est surtout associé à des termes de (non-)déplacement (*percée jusqu'à, être, ne pas être* ou *arriver*) ou à des causatifs de déplacement (*rapporter / ramener qqch à la maison*) : grammaticalement analysable comme un complément locatif, il se construit sémantiquement avec une majorité de prédicats indiquant le retour (*être de retour, revenu, rentre, reviendra, revenir, retour triomphal, rentré...*), suivis quantitativement par des verbes exprimant que l'on ne bouge pas de chez soi (*rester, être*), tous susceptibles d'être suivis de l'expression d'une activité (*J'étais tout le temps à la maison derrière mon ordinateur ; rester à la maison et s'occuper des enfants ; rentrer à la maison le midi m'occuper des enfants ; etc.*).
- Dans le deuxième type de contexte le plus fréquent, *à la maison* s'analyse comme un ajout de verbe ou de syntagme verbal ; il concerne des activités diverses ou des situations entre lesquelles il est difficile d'établir un lien : *dort souvent, être caché, jouer au piano, recettes réalisées, endoctrinés, se radicaliser, jouer, bricoler leurs chansons, travailler, garder sa fille, écrans connectés, recharger leur véhicule, se faire contrôler, recevoir, divertissement...* Bon nombre de ces énoncés ont trait à l'école et à l'activité scolaire : *expériences à faire, l'école, jeux scolaires, suivre le cours, exercices à faire, instruits, dispenser les bases, suivre les cours, la maîtresse, les devoirs...* Ainsi, paradoxalement, *à la maison* se trouve souvent associé à des activités dont on s'attendrait précisément à ce qu'elles se déroulent ailleurs (par exemple, *être endoctriné, se faire contrôler* ou *recharger son véhicule*) ; si « les devoirs à la maison » sont une banalité, les autres contextes sont, là aussi, relativement surprenants, évoquant le propre de l'école plutôt que des activités habituelles à la maison.
- Le troisième type de contexte, où *à la maison* est également ajout de verbe ou de syntagme verbal, évoque un ressenti (*se sentir*), un mode de vie plus conforme à ce que l'on s'imagine être « la vie à la maison » (*vivre en pagne, parler de travail, parler russe, règles mises en place, rapporter des témoignages, appeler quelqu'un de telle manière*).

CONCLUSION SUR LES OBSERVATIONS

À l'issue de cette étude très limitée – car pour être à la hauteur des hypothèses que l'on est en droit d'avancer au vu des observations qu'elle a permises, l'examen devrait couvrir d'autres types de corpus que nous ne pouvons nous permettre ici –, nous avancerons les conclusions provisoires suivantes :

- d'une part, chaque lexème employé en discours est pourvu d'une charge connotative qui n'est pas forcément *a priori* prévisible d'après l'ensemble des traits sémantiques formant sa définition en langue (comme vu à propos de *studio* ou de *villa*),
- d'autre part, chaque lexème employé en discours est associé à un ensemble de représentations qui n'est pas forcément *a priori* prévisible d'après son signifié

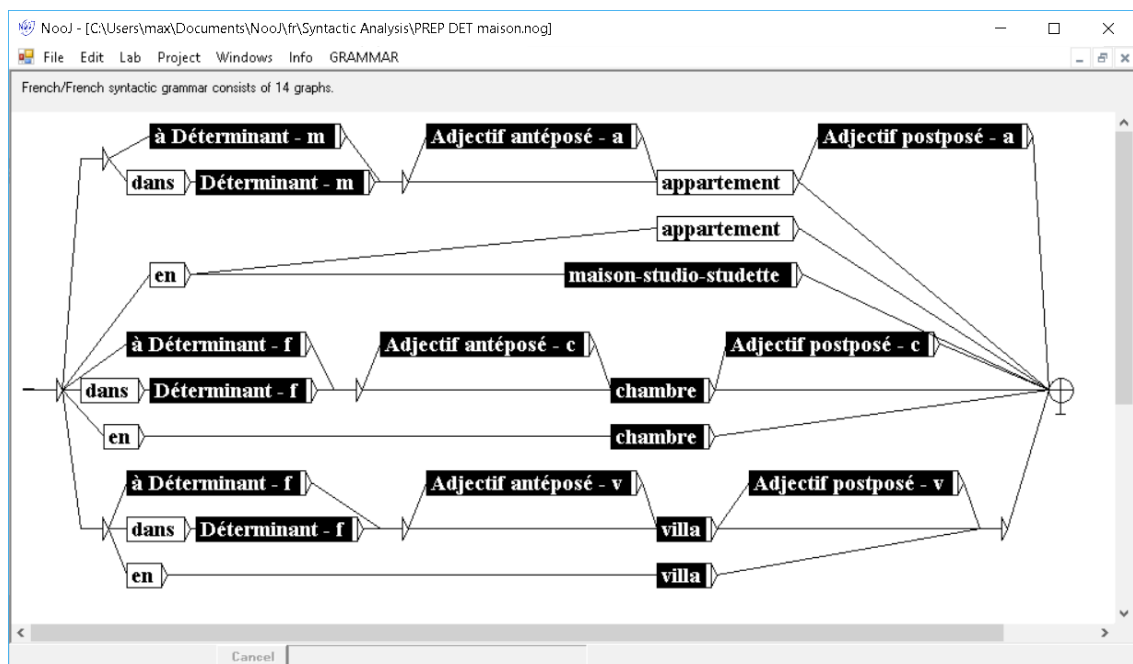
(hypothétique) ; ainsi a-t-on observé que la maison pouvait être conçue comme le lieu d'activités scolaires que l'on aurait spontanément plutôt attribuées à d'autres lieux plus spécialisés ;

– enfin, ces charges connotatives et représentations apparaissent liées à des corpus de genres différents, c'est-à-dire dépendent moins de la taille des données consultées que de leur source, de leur destinataire et de leur objectif communicationnel.

L'ensemble de ces descriptions, calculs et synthèses nous semble nécessaire au test d'une hypothèse, et leur accumulation constituait une démonstration suffisante, dans la seconde moitié du XX^e siècle, pour sa validation. A l'heure actuelle, un formalisme en est une mise à l'épreuve plus rigoureuse, permettant de vérifier si les données ainsi décrites et organisées sont susceptibles de s'organiser en un modèle capable de reconnaître ou de produire des suites grammaticales et acceptables.

IMPLÉMENTATION

Nous avons implémenté une grammaire avec le logiciel NooJ⁷ pour rendre compte du vocabulaire utilisé dans notre corpus, tel qu'analysé *supra*.



Rappelons que ce logiciel permet aussi de représenter des grammaires graphiquement : les expressions reconnues sont celles qui peuvent être épelées en partant du nœud initial (la flèche horizontale à gauche du graphe) et qui rejoignent le nœud terminal (la cible à droite du graphe). Les nœuds sur fond noir représentent des graphes imbriqués. Par exemple, le graphe « Adjectif

⁷ NooJ est une plateforme linguistique utilisée pour formaliser une dizaine de types de phénomènes linguistiques (orthographe, lexique, syntaxe et sémantique), cf. (Silberstein 2015). Ici, nous avons utilisé NooJ comme outil d'exploration de corpus, en reprenant la méthodologie présentée par (Silberstein 2018) : construction de ressources lexicales, étude des concordances correspondantes, formalisation des contextes avec des ressources grammaticales.

antéposé - a » représente l'ensemble des adjectifs qui peuvent apparaître devant le nom *appartement* ; ils sont susceptibles d'être précédés d'un des adverbes *plus, tout, très* ou *trop*.

1. L'objectif de la grammaire

L'objectif est de décrire dans cette grammaire les expressions du journal *Le Monde*, lequel ne comporte pas de séquences telles que *dans une chambre de bonne à loyer modéré* ni *dans une petite villa*, ce qui le distingue d'autres corpus – par exemple l'association *petite villa* se trouve dans les annonces de locations saisonnières⁸. Etablir les grammaires propres à chaque corpus permet précisément une comparaison entre ces derniers : si grands soient-ils, tous ne sont pas également représentatifs des constructions ou des distributions d'une langue.

2. Ce dont rend compte la grammaire

Cette grammaire rend bien sûr compte de certaines interdictions rencontrées dans le corpus ; ainsi, nous n'avons pas gardé la possibilité pour la préposition *en* d'être suivie d'un nom modifié par un adjectif : *en jolie chambre de service, en villa somptueuse*. Certains adjectifs peuvent apparaître avant ou après tous ces noms, par exemple :

*dans un ancien (appartement | studio) ; dans un (appartement | studio) ancien ;
dans une ancienne (chambre | maison | studette | villa)*

D'autres n'apparaissent qu'avec certains d'entre eux :

un appartement à loyer modéré ; une chambre d'hôtel sans eau chaude ; une villa cossue.

Nous avons donc décrit quatre ensembles d'adjectifs pour rendre compte des noms *appartement* (Adjectif – a), *maison* (Adjectif – m), *chambre* (Adjectif – c) et *villa* (Adjectif – v).

En revanche, nous ne nous sommes pas contentés de ne décrire que les séquences qui apparaissent effectivement dans le corpus du journal *Le Monde 2016* :

— nous avons généralisé l'utilisation des adverbes *plus, tout, très* et *trop* à de nombreux adjectifs (sauf, bien entendu, ceux avec lesquels on obtiendrait des suites agrammaticales : * *un appartement très à loyer modéré*) ;

— nous avons vérifié et intégré systématiquement la possibilité de placer les adjectifs avant et/ou après le nom.

3. Conclusion et comparaison

La grammaire ainsi constituée reconnaît (et peut aussi produire) potentiellement plus de 300.000 expressions. Son élaboration a nécessité une analyse linguistique mettant en jeu la connaissance des unités grammaticales (catégories, fonctions, constructions) et lexicales afin, par exemple, de ne sélectionner que les emplois de *studio* ou de *chambre* qui intéressent la recherche. Notons qu'un simple relevé des formes orthographiques⁹ aurait complètement faussé

⁸ Ainsi : « (17) St Denis d'Oléron bord mer petite villa tout confort prix hors saison int./.../ » (*Notre Temps* n°598, p. 151).

⁹ Ce que font les outils statistiques traditionnellement utilisés pour l'analyse de corpus textuels, comme par exemple Hyperbase (cf. Brunet 2010), Iramuteq (cf. Ratinaud 2009), Lexico (cf. Lamalle et alii,

le corpus de qui s'intéresse à l'habitat et à ses représentations sociales – ce qui est le cas du présent article – en y intégrant un grand nombre d'occurrences de la forme *studio* dans les sens de « studio d'enregistrement », « studio de télévision », « studio photo », etc. et de la forme *maison* dans les sens de « maison d'arrêt », « maison blanche », « maison de courtage », etc. De même, la prise en compte des collocations n'aurait pas suffi, puisque par exemple le seul repérage orthographique de séquences comme *chambre froide* masquerait leur ambiguïté (*mettre la viande en chambre froide* vs *grelotter dans la chambre froide*) : seule une grammaire du contexte minutieusement réglée peut en effet distinguer ces séquences (dans ce cas, grâce à la préposition *en*), et, dans les cas difficiles lorsque le contexte immédiat ne suffit pas, rendre compte des ambiguïtés non résolues.

CONCLUSION GÉNÉRALE

Si, « culturellement (ou socio-culturellement), chaque langue (et donc chaque lexique) a sa propre histoire /.../, car elle répercute les expériences du groupe » (Rousseau, 2019 : 31), l'analyse des discours par ceux mêmes qui parlent cette langue ne leur en réserve pas moins quelques surprises.

Ainsi avons-nous considéré comme inattendu le peu d'occurrences consacrées aux logements d'une seule pièce dans le corpus du *Monde*, compte tenu de la population objectivement concernée par ce type d'habitation. De même avons-nous observé avec étonnement les occupations les plus représentées au sein du foyer des Français à travers la distribution à *la maison*.

Nous avons également découvert qu'un type de corpus n'illustre pas forcément les mêmes distributions qu'un autre, de genre différent : en l'occurrence, le discours journalistique du *Monde* donne aussi bien de *studio*, *villa* ou *appartement* une représentation différente de celle de *GOOGLE* – même s'il subsiste un point commun : la connotation négative pour *studio* ou *appartement*, le luxe pour *villa*. Cela s'explique, naturellement, par la fonction et les destinataires de l'information des deux médias, mais il n'en reste pas moins que la leçon à tirer de ce constat est sans doute que, plus que de grands corpus, ce sont des corpus diversifiés qui sont utiles à la description et à l'explication linguistiques.

Références des travaux cités

Brunet E., 2010. HYPERBASE : Manuel de référence. 2010. (hal-01362721).

Courdesses, L., 1971, « Blum et Thorez en mai 1936 : analyses d'énoncés », *Langue française* 9 : 22-33.

Dubois, J., 1962, *Le Vocabulaire politique et social en France de 1669 à 1872 – à travers les oeuvres des écrivains, les revues et les journaux*. Thèse pour le Doctorat ès Lettres présentée à la Faculté des Lettres et Sciences Humaine de l'Université de Paris. Librairie Larousse, Paris.

2002), TXM (cf. Heiden 2010).

- Dubois, J., 1979, « Annexe grammaticale », *Dictionnaire du français langue étrangère II*, Paris, Larousse.
- Dubois, J. & F. Dubois-Charlier, 1997, *Les Verbes français*, Paris, Larousse-Bordas.
- Gross, G., 1994, « Classes d'objets et description des verbes », *Langages* 115 : 15-30.
- Guillaume, G., 1919, *Le problème de l'article et sa solution dans la langue française*, Paris, Nizet & Québec, Presses de l'Université Laval.
- Harris, Z. S., 1954 trad. 1970, « La Structure distributionnelle », *Langages* 20 : 14-34.
- Heiden S., Magué J.-P., Pincemin B. « TXM : Une plateforme logicielle open-source pour la textométrie - conception et développement ». 10th International Conference on the Statistical Analysis of Textual Data - JADT 2010, Jun 2010, Rome, Italie. pp.1021-1032. {halshs-00549779}.
- Lamalle, C., Martinez, W., Fleury, S., Salem, A., Fracchiolla, B., Kuncova, A., & Maisondieu, A., 2002. *Lexico 3, Outils de statistique textuelle*. Université de la Sorbonne Nouvelle.
- Leeman, D., 2017, « Les paradigmes comme échos / véhicules des représentations sociales : l'exemple de structures de type Prép + N. », *Signata - Annales des sémiotiques* 8 :101-128.
- Rousseau, A., 2019, « Réflexion sur l'étude du lexique », *Langages* 214 : 19-32.
- Ratinaud, P., 2009. « IRAMUTEQ : Interface de R pour les analyses multidimensionnelles de textes et de questionnaires ». Computer Software). *Acesso em vol.15*.
- Silberztein, M., 2010, « La formalisation du dictionnaire LVF avec NooJ et ses applications pour l'analyse automatique de corpus ». *Langages* n° 179-180, pp. 221-241.
- Silberztein, M., 2015, *La formalisation des langues : l'approche de NooJ*. ISTE Ed.: Londres. (429 pages).