

LES ADJECTIFS DE FORME *INXABLE* EN FRANÇAIS

Georgette Dal, UMR 8163 « STL », CNRS & Un. Lille 3 et Lille 1
Natalia Grabar, Inserm U729, SPIM,
Stéphanie Lignon, UMR 5610 « ERSS », CNRS & Un. Toulouse-Le Mirail et
Un. Haute-Alsace
Clément Plancq, UMR 7110 « LLF », CNRS & Un. Paris 7
Delphine Tribout, UMR 7110 « LLF », CNRS & Un. Paris 7
François Yvon, GET/ENST et CNRS/LTCI

Introduction

La présente étude s'inscrit dans un travail collectif mené dans le cadre de l'action 1 du GDR 2220 du CNRS « Description et modélisation en morphologie » dirigé par Bernard Fradin. Cette action, pilotée par Georgette Dal et intitulée « Aspects de la productivité morphologique », réunit des linguistes, des informaticiens et des talistes¹. L'un des objectifs qu'elle se donne à terme est de proposer une cartographie de la productivité des principales règles de construction de lexèmes du français, comme il en existe pour d'autres langues, mesurée à partir de relevés en corpus.

Lors de nos premières analyses sur corpus destinées à calculer la productivité de la suffixation en *-able* du français, nous avons été confrontés à un nombre important d'adjectifs de forme *inXable*, et nous avons pu constater *in situ* un fait bien connu: celui des adjectifs en *inXable* sans *Xable* associé. Le fait est en effet périodiquement signalé par les linguistes (Staaff 1928, Peter 1949, Gaatone 1987, Anscombe 1994, Corbin 1997, Apothéloz 2003, Schwarze 2007)², ainsi que par les lexicographes. Il explique l'analyse par la parasyntèse que reçoivent certains adjectifs de forme *inXable* (Peter 1949: 169, Grevisse 1988: 254, Béchade 1992: 134). Le phénomène s'est présenté à nous dans un quart des cas. Voici quelques exemples d'adjectifs en *in-* sans *Xable* associé, extraits de notre corpus de travail constitué de plus de 70% des années 1995 et 1999 du journal *Le Monde*. L'absence constatée est celle de l'adjectif en *Xable* dans notre corpus, indépendamment de son éventuelle attestation

¹ Pour des raisons diverses, plusieurs membres de cette action, pourtant assidus aux réunions de l'opération, n'ont pas été signataires de ce papier. Nous pensons particulièrement à Bernard Fradin, à Nabil Hathout, à Fiammetta Namer et à Pierre Zweigenbaum. Nous tenons à les remercier chaleureusement pour leurs remarques, dont s'est nourri le présent article.

² L'existence de préfixés négatifs sans contrepartie positive a pareillement été soulignée pour d'autres langues, comme l'anglais (Zimmer 1964, Horn 2001).

dans d'autres corpus textuels ou dictionnairiques. Le nombre d'occurrences relevées figure entre crochets:

- (1) *imbattable* [33], *imperturbable* [49], *inaltérable* [27], *inconsolable* [14], *increvable* [15], *indémêlable* [2], *indomptable* [24], *infroissable* [1], *intouchable* [7]

Malgré cela, et de façon apparemment paradoxale, linguistes et lexicographes s'accordent à souligner l'accointance de la préfixation en *in-* pour les bases en *-able* (voir entre autres Zimmer 1964: 49, Gaatone 1987: 81, Anscombe 1994, Apothéoz 2003, *TLFi* s.v. **in-** A.1.), à tel point que, pour Schwarze (2007), seuls les adjectifs en *in-* construits sur des bases en *-able* résultent de l'application d'une règle: les autres seraient formés par imitation analogique.

Dans cette contribution, nous nous donnons pour objectif de confronter une hypothèse théorique faite à propos de la préfixation en *in-* du français avec la réalité de données effectivement produites en corpus, en focalisant notre attention sur les adjectifs de forme *inXable*. Pour mener à bien cette recherche, nous avons conçu un certain nombre d'outils visant à détecter, en corpus, les adjectifs construits en *in-* et en *-able* du français, ainsi qu'une interface permettant de consulter les contextes d'utilisation effective de ces adjectifs. Dans le § 2, nous exposerons notre hypothèse théorique de départ ainsi que les prédictions dont elle s'assortit. Dans le § 3, nous présenterons notre corpus, ainsi que les différents outils que nous avons élaborés pour réaliser notre recherche. Dans le § 4, nous choisirons quelques couples d'adjectifs de forme *inXable* / *Xable* que nous observerons en contexte de façon à évaluer la validité de l'hypothèse théorique de départ.

1. Hypothèse

1.1. Formulation de l'hypothèse

Partant du constat qu'en français, les néologismes adjectivaux en *in-* sont très massivement construits sur des simples en *-able*³, la première hypothèse que nous ferons est que la règle de construction de lexèmes dont l'exposant est le préfixe *in-* (nous appellerons désormais cette règle la RCLⁱⁿ⁻) est essentiellement, voire exclusivement, disponible avec ce type morphologique de simples. En effet, les très nombreux travaux de Baayen sur la productivité reposent sur le fait que, dans un corpus d'une taille suffisamment importante, les *hapax legomena* (types ne possédant qu'une occurrence dans le corpus considéré) tendent à se confondre avec les néologismes, *i.e.* avec les formes de lexèmes non répertoriés dans les dictionnaires (pour une discussion, voir Dal 2003). Pour vérifier la préférence forte de la préfixation par *in-* pour des simples en *-able* dans la période actuelle, nous

³ Dans ce travail, nous préférons parler de 'simple' plutôt que de 'base', *simple* étant à entendre « morphologiquement plus simple que le construit ». Ainsi, *USABLE* est construit relativement à *USER*, mais simple relativement à *INUSABLE*.

avons extrait de notre corpus l'ensemble des adjectivaux hapaxiques de forme *inX*. Nous avons ensuite vérifié si chacun de ces adjectifs était ou non attesté dans la version électronique du *Trésor de la Langue française*. Si la tendance d'un fort recouvrement entre hapax et néologismes se vérifie mal dans le cas qui nous intéresse (sur les 148 hapax en *inX* du *Monde* 1995, 14 seulement ne sont pas attestés dans le *TLF*; les chiffres sont comparables dans *Le monde* 1999, dans lequel 8 seulement des 132 hapax de forme *inX* sont absents du *TLF*), il est intéressant de remarquer que les hapax néologiques sont tous construits sur un simple en *-able* dans le *Monde* 1995 (2a). Dans le *Monde* 1999 (2b), la tendance est moins forte, mais demeure néanmoins:

- (2) a. *imprésentable, incasable, incritiquable, indansable, indéboulonnable, indécidable, indéductible, indémodable, indispersable, infilmable, inintégrable, insituable, insoignable, insortable*
 b. *imbrûlé, incontenable, indécouvert, inenchaînable, inlouable, inmontable, inofficiel, intachable*

Etant admis la préférence actuelle forte de la RCLⁱⁿ⁻ pour des simples en *-able*, notre objectif étant de rendre compte du fonctionnement actuel de la RCLⁱⁿ⁻, la régularité que nous dégagerons ne vaut par conséquent que pour les seuls simples en *-able*, et ne prétend pas rendre compte d'adjectifs comme *incertain, incohérent* ou *injuste*, qui sont entrés depuis longtemps dans le lexique (respectivement en 1329, 1751 et 1293) et dont les bases ne satisfont pas le type morphologique retenu. Elle ne dit rien non plus des adjectifs fortement lexicalisés, fussent-ils construits sur des simples en *-able*, en particulier de ceux qui, comme *incroyable* ou *inouvable*, présentent une valeur superlative. A l'instar d'Apothéoz (2003) auquel nous renvoyons pour plus de détails, nous imputerons cette possibilité à l'application de mécanismes inférentiels à partir du sens construit par la préfixation en *in-*. La RCLⁱⁿ⁻ ne serait donc pour rien dans ces emplois (Fradin 2003: 272-273 fait une hypothèse analogue pour certains *Xable*).

Notre hypothèse consiste à poser qu'en synchronie, la RCLⁱⁿ⁻ forme des lexèmes exprimant la non-satisfaction d'une propriété attendue, l'adjectif en *in-* ne donnant aucune indication quant au degré de non-satisfaction de cette propriété. La propriété attendue, exprimée par le simple en *-able*, vaut pour la catégorie que nomme le nom dont le construit en *in-* prédique une propriété (nous parlerons désormais de nom sémantiquement recteur, voire plus simplement de nom recteur).

Etant admis que, si elle est attendue pour une catégorie d'objets, une propriété donnée ne peut qu'être ou ne pas être satisfaite, même si c'est à des degrés divers, cette formulation revient à considérer qu'à l'instar de ses homologues anglaises en *un-* et en *in-*, la préfixation en *in-* du français forme fondamentalement des contradictoires de leurs lexèmes-bases plutôt que des contraires⁴. Comme en anglais

⁴ Nous utilisons ici la distinction contraires / contradictoires de Horn (2001). En deux mots, contraires et contradictoires partagent la propriété de ne pas pouvoir être simultanément vrais d'un

(voir Horn 2001: 273, qui reprend Jespersen 1917: 144), il peut arriver que les préfixés en *in-* du français soient interprétés comme des contraires de leurs lexèmes-bases. Cependant, selon Jespersen, en anglais échappent à cette tendance précisément les préfixés sur des bases en *-able*, ce que Zimmer (1964) explique en arguant de la productivité de la préfixation en *un-* avec ce type morphologique de bases. L'observation nous semble transposable au français.

1.2. *Prédictions*

Telle qu'elle est formulée, notre hypothèse s'assortit d'un certain nombre de prédictions que nous énoncerons d'abord avant de les justifier.

1.2.1. *Prédiction 1*

Pour un nom sémantiquement recteur donné, la situation normale est que (a) soit le simple en *-able*, (b) soit le construit en *in-* fassent défaut⁵.

Selon la prédiction 1, le cas régulièrement signalé comme déviant dans certains travaux rappelés en introduction (*inXable* sans *Xable* associé) serait finalement la situation normale si l'hypothèse que nous formulons est juste. Pour nous en revanche, le cas parfois implicitement considéré comme banal (tout adjectif en *-able* a un correspondant en *in-*, attesté ou possible⁶) est, lui, discutable. C'est ce cas que cherche à expliquer la prédiction 2.

1.2.2. *Prédiction 2*

Si le construit et le simple sont tous deux attestés, (a) soit les propriétés qu'ils prédisent ne s'appliquent pas aux référents des mêmes noms, (b) soit, si le nom recteur est commun, le référent de ce dernier possède la propriété qu'exprime le simple au-delà ou en deçà de ce qui est attendu pour la catégorie d'objets nommés par lui.

2. **Justification**

2.1. *Prédiction 1*

2.1.1. *Prédiction 1.a*

En dehors des énoncés définitoires, on n'énonce en général pas les propriétés constitutives des objets, selon un principe de pertinence qui interdit de parler de

même objet. Cependant, seuls les contraires autorisent un moyen terme. Horn (2001: 270) oppose ainsi *black* et *nonblack*, qui sont des contradictoires, et *happy* et *sad*, qui sont des contraires.

⁵ On se place ici dans une définition du lexème comme unité de base de la morphologie (Fradin & Kerleroux à paraître), en amont de l'insertion en syntaxe. Ainsi, quand nous disons que le simple fait défaut, nous entendons qu'il fait défaut en tant qu'atome pour la syntaxe. Cela ne l'empêche pas d'exister morphologiquement, en tant que lexème sur lequel s'applique la préfixation en *in-*, indépendamment de sa possible utilisation en syntaxe.

⁶ Leeman (1992: 45) réserve cette possibilité aux *Xable* pouvant être mis en relation avec un verbe (*abaissable*, *abrogeable*, etc.), par opposition à des adjectifs comme *abominable*, *aimable*, dans lesquels la reconnaissance d'un verbe fait davantage problème. Sur le constat que la préfixation en *in-* fait parfois défaut, voir également Anscombe (1994).

chaussures à semelles ou d'homme à ventre, toutes les chaussures ayant des semelles et tous les hommes des ventres. Dès lors, si, conformément à notre hypothèse, on admet que l'adjectif de forme *inXable* exprime la non-satisfaction d'une propriété attendue pour la catégorie d'objets nommée par son nom recteur, on s'attend à ce que le simple en *Xable*, qui nomme la propriété attendue, fasse défaut, à nom recteur constant⁷. Cette première partie de notre prédiction rejoint tout à fait les positions de Leeman & Meleuc (1990: 32-33), de Anscombe (1994: 304 *sqq.*) reprises dans Corbin (1997), ainsi que celle, ancienne, de Staaff (1928) telle que l'expose Peter (1949: 169):

Das Vorhandensein des Negativa ohne Simplex [...] erklärt Staaff mit des folgenden wichtigen Gesetz vor der Oekonomie in der Wortbildung: „La langue ne crée que les mots dont elle a besoin. Et il y a des qualités dont la présence est si normale qu'on n'a presque jamais besoin de la constater“

2.1.2. Prédiction 1.b

Réciproquement, si l'adjectif en *-able* n'exprime pas une propriété attendue pour la catégorie d'objets nommée par son nom recteur, notre hypothèse prédit que la RCLⁱⁿ⁻ ne peut pas s'y appliquer. C'est particulièrement vrai quand il sert à sous-catégoriser la catégorie que nomme son nom recteur (notre observation est en conformité avec Anscombe 1992: 306-07, selon lequel les adjectifs classifiants sont rétifs à la préfixation en *in-*). On citera ici l'adjectif *portable* associé au nom *ordinateur*. Dans *ordinateur portable*, *portable* sert à sous-catégoriser la catégorie des ordinateurs, opérant une partition entre ceux qui peuvent être facilement portés et ceux qui ne le peuvent pas. Pour nommer ces derniers, on parle volontiers d'ordinateurs fixes ou d'ordinateurs de bureau. On pourrait également les nommer *ordinateurs non portables* en recourant à la préfixation par *non-* qui forme elle aussi des contradictoires des lexèmes auxquels elle s'applique, la différence étant qu'il s'agit dans ce cas de contradictoires polaires, par opposition aux contradictoires non nécessairement polaires que forme la RCLⁱⁿ⁻ (aucune propriété attendue ne nous semble en outre en jeu avec la préfixation en *non-*). La dénomination *ordinateurs importables*, où *importable* est construit sur *portable*, nous paraît en revanche exclue pour nommer les ordinateurs de bureau, et les seuls ordinateurs pouvant être dits importables sont des occurrences d'ordinateurs... portables: dans ce cas, la propriété attendue vaut non pas des ordinateurs en général mais de la sous-classe des ordinateurs portables, et l'on retrouve le premier volet de notre hypothèse.

⁷ Comme on va le voir en 2.1.2., nous ne disons pas ici que les adjectifs en *Xable* expriment toujours une propriété attendue. Cela n'est vrai que quand ils servent de base à un préfixé en *in-* (pour des caractérisations récentes de la suffixation en *-able*, voir Fradin 2003 ainsi que Hathout et al. 2003).

2.2. Prédiction 2

2.2.1. Prédiction 2.a

Nous ne nous attarderons pas sur le cas où les noms recteurs sont différents. En effet, une propriété attendue pour les membres d'une catégorie donnée ne l'est pas nécessairement pour ceux d'une autre catégorie. On s'attend donc à ce que simple et construit puissent coexister dans un même corpus avec des noms recteurs différents.

2.2.2. Prédiction 2.b

Le premier volet de la prédiction 1 se situait implicitement dans le cas où le simple exprimait une propriété attendue à un degré standard. En revanche, le principe de pertinence n'est pas enfreint si le simple est employé dans un énoncé à la modalité négative, si la propriété qu'il exprime est possédée à un niveau remarquable par le référent de son nom recteur, ou si elle est d'une quelconque façon modalisée. Cette prédiction suppose par conséquent que le simple exprime plutôt une propriété gradable.

L'exemple du couple *inusable / usable*, cité par Grevisse (1988) et repris par Corbin (1997), illustrera ce qui précède. Les artefacts matériels tels que les rasoirs, les marqueurs, le plancher, les écrans ou les fourreaux de voile sont destinés à s'user. Aussi ne dira-t-on pas d'occurrences particulières de ces catégories qu'elles sont usables, dans la mesure où il s'agit d'une propriété valant a priori de l'ensemble des objets que rassemblent ces catégories, à moins qu'elles ne le soient tout particulièrement (3), particulièrement peu (4), qu'on compare leur propension à l'usure avec celle d'autres membres de leur catégorie (5), qu'on veuille intervenir sur leur usabilité (6) ou, simplement, qu'on leur nie cette propriété (7). Les quelques exemples suivants, relevés sur la Toile le 8 mai 2006, illustrent ces diverses possibilités:

- (3) *Le rasoir est certes très "TRES": – très esthétique – très cher – très irritant – **très usable** – très mauvais quoi*
- (4) *Ce plancher **difficilement usable** constitue un très bon choix pour une famille active*
- (5) *Avec son écran couleur, **moins usable que sur certains autres portables**, avec des sonneries HIFI, mieux que le polyphonique, avec une autonomie importante (5 jours avec appels réguliers), ce téléphone est pour moi un petit bijou*
- (6) *La seule chose qui me préoccupe est de savoir si le fourreau s'use normalement, au bout de quelques mois, ou rapidement, éventuellement comment on pourrait le **rendre moins "usable"***

- (7) *Les marqueurs listés ci-dessus ne sont pas usables, vous pouvez les réutiliser autant de fois que vous le souhaitez*

Autrement dit, le caractère d'étrangeté de l'adjectif *usable*, qui lui vaut d'être rejeté par Grevisse (1988), ne subsiste que si on le considère hors contexte, l'implicite étant qu'il exprime une propriété à un degré standard. Dès lors qu'on l'insère dans un contexte favorisant, l'étrangeté disparaît.

3. Présentation et préparation du matériel linguistique

L'observation des adjectifs de forme *inXable* et *Xable* est faite en corpus, grâce à une interface conçue à cette fin. Nous utilisons huit rubriques thématiques des articles du journal *Le Monde* de l'année 1995, soit 30 599 articles, contenant plus de 15 millions d'occurrences (Tab. 1). Notre approche de détection et d'analyse de ces adjectifs combine les traitements automatiques et l'analyse humaine.

Les outils de traitement automatique des langues permettent d'assurer une analyse systématique et méthodique de gros volumes de données. Nous les utilisons pour le nettoyage et le formatage des corpus, pour leur étiquetage morphosyntaxique et la détection de lexèmes construits, ainsi que pour un premier dégrossissage morphologique. Nous avons de la sorte effectué un nettoyage du format de *Le Monde* d'origine et ventilé les articles en fonction des rubriques. Le tableau 1 présente les huit sous-corpus thématiques ainsi obtenus. La première colonne indique les rubriques étudiées, les suivantes indiquent le nombre d'articles (*Nb Art*) et d'occurrences (*occ*) par rubrique.

Rubrique	Nb Art	occ
Agenda	1 213	605 432
Événements culturels	4 242	2 154 164
France	6 331	3 870 389
International	9 276	3 661 400
Livres	1 949	1 624 924
Programme TV et radio	1 217	883 060
Société	4 009	2 020 013
Événements sportifs	2 362	1 177 669
TOTAL	30 599	15 997 051

Tab. 1. Taille des corpus par rubrique de *Le Monde* en 1995

Les principaux efforts du traitement automatique sont concentrés autour de l'étiquetage morphosyntaxique, dans la mesure où on trouve à ce niveau des informations fondamentales pour l'analyse morphologique. Pour ce faire, nous avons utilisé l'étiqueteur *TreeTagger* (Schmid 1994). Afin d'améliorer la qualité de l'étiquetage, nous avons effectué un préétiquetage des corpus avec un lexique généraliste du français téléchargeable à partir du site www.lexique.org. Les sorties de *TreeTagger* ont ensuite été traitées avec le lemmatiseur du français *Flemm* (Namer 2000), qui permet de corriger certaines erreurs et de calculer les lexèmes qui ne font pas partie des dictionnaires de *TreeTagger*. Finalement, nous avons

appliqué l'analyseur morphologique *Derif* (Namer 2002) qui combine étiquetage morphosyntaxique, règles de construction de lexèmes et listes d'exceptions pour détecter les lexèmes construits et en proposer une analyse morphologique. Malgré le passage au crible de ces divers outils automatiques, les résultats demeurent encore insuffisamment précis et fiables (Evert & Lüdeling 2001, Grabar et al. 2006). Aussi avons-nous terminé la chaîne de traitement par une validation manuelle des propositions issues de l'analyse automatique.

Cette validation manuelle vise à éliminer un certain nombre de lexèmes considérés comme parasites, étant donné l'objectif assigné à l'analyse. Dans le cas qui nous intéresse, nous avons systématiquement éliminé notamment:

(i) les lexèmes dans lesquels l'affixation étudiée n'était pas la dernière opération constructionnelle. Cela a été le cas de l'ensemble des *inXable* quand il s'est agi d'étudier la suffixation par *-able*, dans la mesure où ils résultent de la préfixation en *in-* de simples en *-able*,

(ii) les lexèmes difficilement analysables comme construits en français, par exemple *affable* et *friable* relativement à la suffixation en *-able*,

(iii) les lexèmes comportant une suite graphique accidentellement identique à celui de l'exposant de la RCL étudiée. Ont ainsi été éliminés *faible*, *double*, dont la finale graphique n'est qu'accidentellement superposable aux allomorphes *-ible* et *-uble* du suffixe *-able*,

(iv) les erreurs et fautes d'orthographe: par exemple *constestable*, *amable*, *comprable*.

Selon les cas, la réduction du nombre de lexèmes conservés à l'issue de la validation manuelle peut aller jusqu'à 50 %. Notre matériel de travail est donc constitué des adjectifs en *inXable* et *Xable* validés. Pour leur observation en corpus, nous avons conçu une interface de consultation de leurs contextes, encore appelée « concordancier ».

Le calcul de concordances, qui est un cas particulier de consultation de contextes, se définit par trois paramètres (Lebart & Salem 1994): le pivot (*i.e.* l'unité linguistique dont on souhaite étudier le contexte), la taille du contexte et le tri. Dans notre travail, le choix du concordancier a été dicté par les spécificités et les contraintes suivantes:

– les pivots étaient connus à l'avance: il s'agissait bien sûr des adjectifs en *inXable* et *Xable* conservés à l'issue de la validation manuelle,

– comme les chercheurs participant au présent travail étaient éparpillés sur le territoire français, l'outil de consultation devait être accessible à distance,

– les adjectifs étudiés devaient être visualisables dans une fenêtre suffisamment grande, qui prenne en compte l'ensemble des dépendances syntaxiques autour des noms pour les référents desquels les adjectifs observés prédisaient une propriété, indépendamment de la complexité syntaxique des phrases.

Le premier point présentait un avantage, car il a permis de générer les concordances à l'avance, étant donné un ou plusieurs corpus. C'est également à l'avance que la fonctionnalité du tri a été appliquée. Le deuxième point a nécessité la création d'une interface Web. La dernière contrainte a été plus difficile à

satisfaire. En effet, dans les concordanciers classiques, la largeur des contextes droit et gauche correspond à la largeur de la ligne et se trouve définie par le nombre de caractères. Or, pour l'analyse des contextes, il nous a paru nécessaire de prendre en compte minimalement une unité comme la phrase où, par *phrase*, nous entendons « suite de caractères comprise entre deux ponctuations fortes (point, point d'interrogation, point d'exclamation) », autrement dit une unité dont la longueur est par définition variable. Cette exigence allait donc au-delà des possibilités habituellement proposées par les concordanciers. Il nous a donc fallu choisir entre la préservation de la largeur de la fenêtre graphique lors de la visualisation, et la complétude du contexte syntaxique du pivot (Pincemin et al. 2006). Pour respecter les exigences de notre étude, nous avons choisi la deuxième solution, et nous avons décidé de présenter les adjectifs au sein des phrases dans lesquelles ils apparaissent. Les adjectifs étudiés ont été de surcroît affichés en rouge afin de faciliter leur repérage.

Notre concordancier se présente ainsi sous la forme d'un ensemble de pages html permettant la navigation dans les divers emplois des adjectifs retenus à l'issue de la validation manuelle. La page racine affiche la liste des couples d'adjectifs *Xable / inXable* ordonnés selon leur fréquence dans le corpus, et présente un lien vers les contextes particuliers de chaque couple d'adjectifs.

Pour chaque couple d'adjectifs *Xable / inXable* attesté dans le corpus sont affichées sur une même page toutes les phrases dans lesquelles ils figurent. Les contextes du simple en *-able* et du construit en *in-* sont néanmoins présentés séparément afin de toujours savoir de quel type d'adjectif il s'agit.

La figure 1 donne un aperçu de l'interface de visualisation des contextes des adjectifs *impubliable* et *publiable*:

9 Pourquoi le rémunère -t-il pour ses notes sur Garcia Lorca , Marinetti ou Maïakovski alors qu' elles sont , politiquement , **impubliables** dans un journal aux sympathies salazaristes ?

10 Comme toute l' interprétation de Markowicz , ce parti pris est très controversé dans le milieu de la traduction , qui considère que ce nouveau Dostoïevski est **impubliable** .

11 Il parle en termes **impubliables** de la violence , de la drogue et de l' inceste .

[Contextes inXable \(7\)](#)
[Contextes Xable \(6\)](#)
[Contextes négatifs \(0\)](#) 12 C' est le travail de la postérité : éditions de textes mineurs curiosités , documents , accompagnées de recherches biographiques et de travaux critiques sur des inédits jugés **impubliables** .

13 Mais , à le lire , il est clair qu' il était , alors même censuré , même autocensuré , **impubliable** .

Contextes Xable (6)

1 Ceux-ci , tenus secrets , contenant de l' aveu de son auteur des épisodes non **publiables** (et non exposés !

3 LE DERNIER SONDAGE **publiable** dans la presse écrite avant la période d' interdiction a été réalisé par l' institut BVA pour Le Monde , France 2 et Europe 1 .

4 En souvenir du premier tour de l' élection présidentielle de 1988 , une partie de son entourage , elle , tremblait : les derniers sondages **publiables** avaient donné autour de 24% des intentions de vote au maire de Paris , qui , finalement , était sorti des urnes à moins de 20% des suffrages exprimés .

correspond un simple dans notre corpus. Parmi ces derniers, nous avons opéré une sélection et nous avons systématiquement dépouillé les contextes des couples d'adjectifs dont le rapport entre construit et simple était compris entre 0,5 et 1,5, en rejetant les couples de formes dont l'une au moins présentait un nombre d'occurrences supérieur à 100: les lexèmes très fréquents sont en effet fortement lexicalisés, et il peut arriver que des paramètres, qui n'ont rien à voir avec la règle de construction de lexèmes dont ils sont les sorties, induisent un effet de brouillage sur la règle elle-même. A ce titre, ils peuvent déroger à leur sens construit par la règle qui les a produits.

Les tableaux 2 à 4 recensent ces couples d'adjectifs dont nous avons ainsi étudié les contextes. Nous y donnons les formes relevées, le nombre d'occurrences observées dans l'année 1995 du journal *le Monde*, ainsi que le rapport obtenu entre le nombre de formes en *inXable* et en *Xable* (nous traitons ici *-ible* comme un allomorphe de *-able*):

	<i>inXable</i>	<i>Xable</i>	Rapport <i>inXable/Xable</i> >1
imaginable	69	45	1,53
avouable	28	19	1,47
praticable	27	22	1,22
défendable	19	16	1,18
joignable	8	6	1,33
transportable	7	6	1,16
publiable	7	6	1,16
décelable	6	5	1,2
discernable	5	4	1,25

Tab. 2. Rapport *inXable* / *Xable* >1

	<i>inXable</i>	<i>Xable</i>	Rapport <i>inXable/Xable</i> <1
flexible	54	72	0,75
éligible	40	63	0,63
audible	25	27	0,92
palpable	23	27	0,85
périssable	9	21	0,42
communicable	5	6	0,83
cessible	5	6	0,83

Tab. 3. Rapport *inXable* / *Xable* <1

	<i>inXable</i>	<i>Xable</i>	Rapport <i>inXable/Xable</i> = 1
Jouable	5	5	1
Opérable	1	1	1
formulable	1	1	1

Tab. 4. Rapport *inXable* / *Xable* = 1

4.2. Résultats en contexte

La prédiction 2 formulée dans le § 2.2. prédit que les couples d'adjectifs en *inXable* / *Xable* figurant dans les tableaux 2 à 4 ci-dessous ne sont pas employés avec les mêmes noms recteurs (prédiction 2.a), ou, s'ils le sont, que la propriété attendue exprimée par le simple ne l'est pas à un degré standard (prédiction 2.b). C'est ce que nous allons vérifier en contexte en choisissant quelques couples représentatifs.

4.2.1. *Impubliable* / *publiable*

Dans l'année 1995 du journal *Le Monde*, *impubliable* compte sept occurrences, et *publiable* six, dont une précédée de *non-* (la faible lexicalisation de ce couple se vérifie dans *Le Monde* 1997, qui compte deux occurrences de chaque adjectif). Lorsque l'on étudie ces occurrences en corpus, on s'aperçoit que ces adjectifs ne sont pas employés avec le même type de noms recteurs:

(i) ceux d'*impubliable* réfèrent à des types d'écrits (*roman, notes, inédits*), à leurs auteurs (*nouveau Dostoïevski*) ou à des propos tenus (*termes*),

(ii) ceux de *publiable* sont (a) *vie personnelle*, (b) *sondage / enquête d'opinion*. En première analyse, la répartition (i) / (ii) est conforme à notre prédiction 2.a. (types de noms recteurs différents) et, par ricochet, à la prédiction 1. En effet, il est attendu, pour des écrits ou leurs auteurs, de pouvoir être publiés. Il est donc prévisible que *publiable* ne soit pas utilisé pour prédiquer une propriété d'écrits ou d'auteurs, mais, qu'en revanche, la RCLⁱⁿ⁻ puisse s'y appliquer. De fait, c'est bien la non-satisfaction de la propriété attendue 'pouvoir être publié' qu'exprime *impubliable* dans les exemples (8)-(10) issus du *Monde* 1995:

- (8) *avant de [...] concevoir un roman qu'il jugea **impubliable***
- (9) *Pourquoi le rémunère-t-il pour ses notes sur Garcia Lorca, Marinetti ou Maïakovski alors qu'elles sont, politiquement, **impubliables** dans un journal aux sympathies salazaristes ?*
- (10) *ce parti pris est très controversé dans le milieu de la traduction, qui considère que ce nouveau Dostoïevski est **impubliable***

En revanche, 'pouvoir être publié' n'est pas une propriété attendue des vies personnelles. Il peut néanmoins être pertinent de rendre publique une vie personnelle, si bien que rien ne s'oppose à ce que *publiable* puisse être régi par *vie personnelle*:

- (11) *Parler de la façon dont j'élève mes enfants relève du domaine de la vie personnelle **publiable** parce que c'est moi sans être vraiment moi*

Les emplois que nous avons relevés de *publiable* avec un nom recteur référant à un sondage ou à une enquête d'opinion demandent, eux, une analyse qui aille au-delà du simple nom recteur. En effet, les sondages et autres enquêtes d'opinion sont

normalement destinés à être rendus publics sous forme écrite. *Publiable* exprime donc une propriété attendue de ce type d'entités. Sont cependant interdits de publication les sondages ou enquêtes d'opinion ayant lieu quelques jours avant les élections présidentielles. Dès lors, on fait une partition, parmi les sondages politiques, entre ceux qui sont publiables et ceux qui ne le sont pas, et *publiable* fonctionne comme sous-catégorisateur. Or, c'est précisément de ce type de sondages pré-électorales qu'il s'agit pour quatre des six occurrences de *publiable* que nous avons relevées dans notre corpus (on notera la présence systématique de l'adjectif *dernier*, antéposé au nom recteur de l'adjectif en *-able*):

- (12) *Le dernier sondage **publiable** dans la presse écrite avant la période d'interdiction*
- (13) *l'événement redevient le traditionnel face-à-face qui oppose les deux candidats du second tour devant des millions d'électeurs, dont, selon les dernières enquêtes d'opinion **publiables**, 20% n'ont pas encore choisi celui dont ils veulent faire le président de la République*
- (14) *Le dernier sondage **publiable**, réalisé pour Le Monde par BVA auprès de 1006 personnes, vendredi 28 avril, confirme l'avance de Jacques Chirac sur Lionel Jospin dans les intentions de vote*
- (15) *[...] les derniers sondages **publiables** avaient donné autour de 24% des intentions de vote au maire de Paris [...]*

S'il fallait qualifier les sondages réalisés pendant la semaine précédant les présidentielles, on parlerait de sondage non-publiable.

Nous n'avons relevé aucune occurrence de *sondage impubliable* dans notre corpus. La possibilité existe cependant, comme le montre l'exemple suivant relevé sur la Toile le 26 juillet 2006:

- (16) *Pour le meilleur et pour le pire, les œuvres interdites, les rapports mis au placard, les informations que l'on tente de cacher, les sondages **impubliables** seront en ligne. La transparence concernera les marchés comme le pouvoir*

Si les sondages de (16) ne peuvent pas être publiés, ce n'est pas parce que la législation s'y oppose, mais parce que leur contenu est sulfureux. Autrement dit, dans l'adjectif *impubliable* de (16), le simple n'a plus la valeur classifiante qui est la sienne dans les exemples (12) à (15).

4.2.2. *Impraticable / praticable*

Impraticable et *praticable* comptent respectivement vingt-sept et vingt-deux occurrences dans notre corpus, soit un rapport de 1,22 (à titre de comparaison, le rapport est strictement le même dans *le Monde* 1997, mais avec respectivement 65 et 53 occurrences). Contrairement au cas précédent, ils partagent la plupart de leurs noms recteurs. Ces derniers réfèrent soit à des voies de communication (*route, piste, rue, venelle*, etc.), soit à ce que, faute de mieux, nous appelons des dispositifs mentaux ou physiques (*idée, procédé, procédure, proposition, loi*, etc.).

S'agissant des voies de communication, on s'attend à ce qu'elles soient praticables, autrement dit utilisables par les usagers (piétons, automobilistes, etc.). Dire d'une voie de communication qu'elle est praticable constitue donc une tautologie, du moins si elle possède cette propriété à un degré standard, et, selon notre prédiction 1, seul l'adjectif en *in-* devrait être attesté, comme dans l'exemple (17):

(17) *La moitié des rues et des venelles de l'enclave sont **impraticables***

L'emploi, récurrent, de *praticable* avec un nom référant à une voie de communication demande donc explication. Or, un examen précis du contexte d'utilisation de cet adjectif dans notre corpus montre que, conformément à notre prédiction 2.b, *praticable* soit n'exprime pas une propriété attendue à un degré standard (18), soit que les voies qui possèdent cette propriété attendue sont les seules à avoir conservé cette propriété ((19) et (20)), si bien que, dans un contexte restrictif, il devient pertinent de mentionner la possession de cette propriété:

(18) *La piste est sinueuse, parfois difficilement **praticable***

(19) *Les embouteillages monstrueux sur les seules routes **praticables***

(20) *une seule route le liant à Rennes restant **praticable***

S'agissant des dispositifs mentaux ou physiques, l'attente est la même que pour les voies de communication: un procédé ou une procédure, par exemple, doivent par définition être praticables, i.e. pouvoir être mis à exécution. C'est bien la perte – annoncée – de cette propriété attendue qu'exprime *impraticable* en (21):

(21) *au bénéfice d'une procédure protectrice des droits de la victime tellement complexe qu'elle finirait par être **impraticable***

Quant à *praticable*, de nouveau, il est remarquable que, dans notre corpus, il n'apparaisse qu'avec des noms dont les référents ne possèdent pas la propriété qu'il exprime à un degré standard (22), ou pour lesquels on déplore la rareté de cette propriété attendue (23):

- (22) [...] *cette dernière idée semble, en réalité, peu **praticable***
- (23) [...] *il constate que l'expérimentation est rarement **praticable** et que les faits sont d'une extrême complexité*

S'agissant du couple *praticable / impraticable*, les données que nous fournit l'année 1995 du journal *le Monde* sont donc conformes à notre prédiction 2.b.

4.2.3. *Injoignable / joignable*

Le troisième couple que nous examinerons est *injoignable / joignable*, dont les membres comptent dans notre corpus respectivement huit et six occurrences, soit un rapport de 1,33 (dans *Le Monde* 1997, le rapport est de 1,04, avec 22 vs. 21 occurrences).

De façon plus remarquable encore que dans le cas précédent, ils sont utilisés avec le même type de noms recteurs: tous réfèrent à des individus. Si notre hypothèse est valide, le fait qu'*injoignable* prédique de façon récurrente une propriété valant d'individus indique que la propriété 'pouvoir être joint' est une de leurs propriétés attendues. De fait, les contextes d'apparition d'*injoignable* montrent de façon quasi-systématique la perte de cette propriété attendue, comme l'indiquent les structures prédicatives *devenir / se rendre injoignable*, dans lesquelles l'adjectif exprime une propriété acquise:

- (24) [...] *des amis devenus **injoignables***
- (25) *Joël joue beaucoup sur la nature de son travail pour se rendre **injoignable** et garde secret son numéro de téléphone*
- (26) *Puis, au moment où l'on se demande si le chanteur n'est pas devenu l'homme **injoignable**, le rendez-vous est pris*

En revanche, les contextes dans lesquels on trouve *joignable* font apparaître que la propriété exprimée par cet adjectif est attribuée au-delà du degré standard (voir la mention « à tout moment (à tout instant) », présente dans les trois exemples). Ceci est conforme à notre prédiction 2.b:

- (27) *Je suis **joignable** à tout moment en gardant ma liberté de mouvement*
- (28) *Il est **joignable** à tout instant*
- (29) *Qui sont **joignables** à tout moment, en n'importe quel point du globe, grâce à la messagerie électronique, mais sans que personne ne sache où ils se trouvent*

Par un effet de corpus, nous ne trouvons aucun *difficilement joignable*, mais la Toile nous en fournit plusieurs (voir les exemples (30) à (32) relevés le 24 juillet 2006). L'explication serait la même que pour les exemples précédents, si ce n'est que la propriété est attribuée ici en deçà du degré standard:

- (30) *Je serai en vacances du 17/07 au 06/08, et serai **difficilement joignable** durant cette période*
- (31) *Je n'ai pas pu avoir les coordonnées de l'assurance du propriétaire du dessus qui est **difficilement joignable***
- (32) *Il est important de préciser ici que ces entreprises sont **difficilement joignables** car elles exercent leur activité à l'extérieur*

4.2.4. *Inéligible / éligible*

Pour finir, nous nous proposons d'examiner en contexte le couple *inéligible / éligible*, dont le rapport de fréquence d'emploi, dans notre corpus, est de 0,63, avec un nombre d'occurrences de, respectivement, quarante et soixante trois. Si nous avons mené notre étude dans *Le Monde* 1997, le rapport aurait été comparable (0,67), mais le nombre d'occurrences (137 vs. 202) aurait dépassé les limites que nous nous sommes fixées (nombre d'occurrences inférieur à 100).

Ce couple est intéressant à plusieurs égards: il pose le problème de la représentativité des corpus, et montre l'existence d'une zone de recouvrement possible entre les préfixations en *in-* et en *non-*, en particulier pour les lexèmes fortement lexicalisés.

Dans *Le Monde* 1995, seules des personnes sont dites *inéligibles*, qu'il s'agisse d'hommes d'affaires, d'hommes ayant une fonction politique (députés, maires, etc.), ou d'individus nommés par leur anthroponyme. Les référents dont *éligible* prédique une propriété sont plus variés: il peut également s'agir de personnes, mais aussi d'étendues territoriales ou administratives, d'actions ou de projets. A douze reprises, *éligible* figure dans le syntagme *en position éligible*.

Sur la Toile et dans d'autres corpus textuels sont en outre dits *inéligibles* (à une mesure gouvernementale, à une aide publique, etc.) diverses zones territoriales (terres, communes...), des projets, des dépenses, etc. Nous renvoyons aux exemples (33) à (36), relevés sur la Toile le 27 juillet 2006:

- (33) *Une terre **inéligible** en 1991 mais mise en culture depuis, peut-elle devenir éligible ? Oui, on ne tient compte désormais que de la situation en 2003. Si la parcelle était en culture en 2003, elle devient éligible aux aides découplées, recouplées et au gel*
- (34) *Ce qu'il faut lire, c'est donc en réalité que d'ici 2007 France Telecom collera du Wimax dans les communes **inéligibles** à l'ADSL*

- (35) *Les règles d'éthique doivent être prises en considération, car leur non-respect peut rendre un projet **inéligible***
- (36) *On peut citer, parmi les dépenses **inéligibles**, les postes de charges suivants: - les frais financiers (...)*

De même, *Le Monde* 1997 et 1998 fournissent chacun deux occurrences de *position inéligible*, contre quatre de *position non-éligible*. Par exemple dans *Le Monde* 1997:

- (37) *Faisant contre mauvaise fortune bon coeur, les deux candidats recalés ont accepté de figurer en position **inéligible** sur la liste,*
- (38) *Devenu sénateur en septembre 1986, il figurait en deuxième position **non éligible** sur la liste conduite par Michel Rocard aux sénatoriales de 1995.*

La conclusion que nous pourrions tirer de la seule observation de l'année 1995 du journal *Le Monde* – *inéligible* apparaît avec une moins grande variété de noms recteurs qu'*éligible* – n'est donc pas nécessairement exportable sur d'autres corpus. Tout au plus peut-on dire que, tendanciellement, *inéligible* exprime une propriété d'individus, mais que d'autres possibilités existent.

Par ailleurs, l'explication que l'on a invoquée dans les cas précédents pour expliquer que le simple en *-able* et le construit en *in-* puissent apparaître avec les mêmes noms recteurs n'est pas recevable ici. En effet, selon notre prédiction 2.b, le simple et le construit sont possibles avec les mêmes noms quand le référent de ces derniers possède la propriété qu'exprime le simple au-delà ou en deçà de ce qui est attendu. Dans le cas d'*éligible*, la propriété qu'il exprime est binaire: une entité donnée est éligible à telle fonction, telle aide si elle satisfait des critères précis, dans le cas contraire, elle ne l'est pas. Elle ne peut donc l'être ni particulièrement, ni particulièrement peu. Par ailleurs, nulle catégorie d'objets n'est définitoirement éligible, il n'y a aucune attente en la matière. Par exemple en (33), 'être éligible' (à une aide publique) n'est pas une propriété attendue des terres agricoles, et ces dernières se répartissent entre celles qui sont éligibles à une aide parce qu'elles satisfont les critères pour en bénéficier, et celles qui ne le sont pas, pour les raisons inverses. Autrement dit, selon notre hypothèse, la RCL^{int} ne devrait pas pouvoir s'appliquer à *éligible*, et, pour former le contradictoire de ce terme, on devrait recourir à la préfixation en *non-* (voir la prédication 1.b.).

On en conclut provisoirement dans le présent article qu'au moins pour les lexèmes fortement lexicalisés (sur la Toile au 10 juin 2006, *inéligible* compte 181 600 occurrences, contre 460 200 pour *non éligible*), il existe une zone de recouvrement possible entre la préfixation par *in-* et celle par *non-*. Cette collision s'observe principalement avec des simples exprimant des propriétés non gradables (elle existe pareillement avec les adjectifs *non constructible / inconstructible*, *non recevable / irrecevable*). Du point de vue de la règle qui les a formés, l'adjectif en *in-* et l'adjectif en *non-* ne sont pas synonymes. Du point de vue du locuteur, la

différence existant entre les deux peut être neutralisée. Seule une étude fine sur corpus contrastant les adjectifs en *in-* et les adjectifs en *non-* pourrait nous en apprendre davantage sur les conditions dans lesquelles se recouvrement s'observe. Elle pourra faire l'objet d'une recherche ultérieure.

Conclusion

Au terme de cette investigation, nous pouvons conclure que les prédictions auxquelles donne lieu notre hypothèse, donc notre hypothèse elle-même, se vérifient plutôt bien en corpus avec des formes peu lexicalisées. En effet, l'hypothèse selon laquelle les lexèmes issus de l'application de la RCLⁱⁿ⁻ à des simples en *-able* expriment la non-satisfaction d'une propriété attendue a été faite sur la règle elle-même, indépendamment des phénomènes de lexicalisation susceptibles d'intervenir par la suite.

Notre approche soulève bien évidemment les problèmes liés à l'utilisation de corpus (représentativité, taille, etc.). Toutefois, les résultats prévus par nos hypothèses se confirment, même si, comme très souvent en morphologie, il ne s'agit que de tendances.

Plus généralement, notre recherche a fait apparaître l'intérêt de combiner théorie et outils d'investigation de corpus. En effet, si nos résultats ne sont somme toute pas très originaux (ils ne sont pas très éloignés de ceux de Leeman & Meleuc 1990 ou de Anscombe 1994), l'analyse que nous avons menée présente à nos yeux l'avantage de se fonder sur des énoncés authentiques, et de dépasser la simple intuition du linguiste. Ainsi, là où Leeman & Meleuc (1990: 33) remarquent que « les adjectifs [en *-able*] forgés [sont] plus facilement acceptables associés à une modalité interrogative [...], à des adverbes comme *facilement* ou *difficilement* », appuyant cette observation d'exemples forgés pour l'occasion (par ex. « Ce mur est facilement abattable ») qui peuvent sembler artificiels, une étude sur corpus donne un poids supplémentaire à cette intuition, au demeurant tout à fait juste.

Références

- Anscombe, Jean-Claude. 1994. "L'insoutenable légèreté morphologique du préfixe négatif *in-* dans la formation d'adjectifs". *La négation: Actes du colloque de Paris X-Nanterre, 12-13-14 nov. 1992*, ed. par Pierre Attal, 299-321. (n° spécial LINX 29). Paris: Centre de Recherches Linguistiques, Université de Paris X-Nanterre.
- Apothéloz, Denis. 2003. "Le rôle de l'iconicité constructionnelle dans le fonctionnement du préfixe négatif *in-*". *Cahiers de Linguistique Analogique* 1. 35-63.
- Béchade, Henri. 1992. *Phonétique et morphologie du français moderne et contemporain*. Paris: Presses Universitaires de France
- Corbin, Danièle. 1997. "Entre les mots possibles et les mots existants: les unités à faible probabilité d'actualisation". *Mots possibles et mots existants: Actes du colloque de Villeneuve d'Ascq, 28-29 avril 1997 (Silexicales 1)*, ed. par Danièle

- Corbin, Bernard Fradin, Benoit Habert, Françoise Kerleroux & Marc Plénat, 78-89. Lille: Université de Lille III.
- Dal, Georgette. 2003. "Productivité morphologique: définitions et notions connexes". *Langue Française* 140. 3-23.
- Evert, Stefan & Anke Lüdeling. 2001. "Measuring morphological productivity: Is automatic preprocessing sufficient?" *Proceedings of the Corpus Linguistics 2001 conference*, ed. by Paul E. Rayson, Andrew Wilson, Tony McEnery, Andrew Hardie & Shereen Khoja, 167-175. Lancaster: Lancaster University.
- Fradin, Bernard. 2003. *Nouvelles approches en morphologie*. Paris: PUF
- Fradin, Bernard & Françoise Kerleroux. à paraître. "L'identité lexématique". *Aperçus de morphologie du français*. ed. par Bernard Fradin, Françoise Kerleroux & Marc Plénat. Paris: Presses Universitaires de Vincennes.
- Gaatone, David. 1987. "Les préfixes négatifs avec les adjectifs et les noms verbaux". *Cahiers de Lexicologie* 50:1. 79-90.
- Grabar, Natalia, Georgette, Dal, Bernard, Fradin, Nabil, Hathout, Stéphanie, Lignon, Fiammetta, Namer, Clément, Plancq, Delphine, Tribout, François, Yvon & Pierre, Zweigenbaum. 2006. "Productivité quantitative de la suffixation par *-Able* dans un corpus journalistique du français". *Actes JADT 2006*, ed. par Jean-Marie Viprey, 473-485. Besançon: Presses Universitaires de Franche Comté.
- Grevisse, Maurice. 1988¹². *Le bon usage. Grammaire française*. Paris / Gembloux: Duculot.
- Hathout, Nabil, Marc Plénat & Ludovic Tanguy. 2003. "Enquête sur les dérivés en *-able*". *Cahiers de Grammaire* 28. 49-90.
- Horn, Laurence R. 2001. *A Natural History of Negation*. Stanford: CSLI.
- Jespersen, Otto. 1917. *Negation in English and other languages*. Copenhagen: Host.
- Lebart, Ludovic & André Salem. 1994. *Statistique textuelle*. Paris: Dunot.
- Leeman, Danielle. 1992. "Deux classes d'adjectifs en *-ble*". *Langue Française* 96. 44-64.
- Leeman, Danielle & Serge Meleuc. 1990. "Verbes en tables et adjectifs en *-able*". *Langue Française* 87. 30-51.
- Namer, Fiammetta. 2000. "FLEMM: un analyseur flexionnel du français à base de règles". *T.A.L.* 41:2.
- Namer, Fiammetta. 2002. "Acquisition automatique de sens à partir d'opérations morphologiques en français: étude de cas". *Actes TALN 2002, Nancy, 24-27 juin 2002* ed. par Jean-Marie Pierrel, 235-244. Le Chesnay: INRIA.
- Peter, Max. 1949. *Über einige negative Präfixe im modernfranzösischen als Ausdrucksmittel für die Gegensatzbildung*. Bern: Francke.
- Pincemin, Bénédicte; Fabrice, Isaac; Marc, Chanove & Michel Mathieu-Colas. 2006. "Concordanciers: thème et variations". *Actes JADT 2006*, ed. par Jean-Marie Viprey, 773-783. Besançon: Presses Universitaires de Franche Comté.
- Schmid, Helmut. 1994. "Probabilistic Part-of-Speech Tagging Using Decision Trees". *Proceedings of the International Conference on New Methods in*

- Language Processing*, ed. par Daniel Jones & Harold Somers, 44-49. Manchester: University of Manchester.
- Schwarze, Christoph. 2007. "La notion de règle morphologique et les mots complexes non construits". *Morphologie à Toulouse : Actes du colloque international de morphologie 4èmes Décembrettes (LSTL 37)*, ed. par Nabil Hathout & Fabio Montermini, . Munich: Lincom Europa.
- Staaff, Erik. 1928. "Etude sur les mots composés avec le préfixe négatif *in-* en français". *Studia Neophilologica* 1. 45-73.
- Zimmer, Karl E. 1964. *Affixal negation in English and other languages: an investigation of restricted productivity*. Supplement to *Word* 20/2, Monograph 5. New York.

Résumé

L'objectif de cet article est double: d'une part, formuler une hypothèse théorique à propos de la préfixation par *in-* du français; d'autre part, confronter cette hypothèse ainsi que les prédictions dont elle s'assortit avec la réalité de données effectivement produites en corpus, en focalisant notre analyse sur les adjectifs de forme *inXable* du français. Notre hypothèse est qu'en français, le préfixe *in-* est l'exposant d'une règle de construction de lexèmes qui forme des lexèmes exprimant la non-satisfaction d'une propriété attendue, cette dernière était exprimée par les lexèmes que cette règle prend en entrée. Notre corpus est constitué de l'année 1995 du journal *Le Monde*. L'étude est appuyée par d'autres années du *Monde* et par la consultation de la Toile. Pour mener à bien notre étude sur corpus, nous combinons outils automatiques, dont certains ont été conçus à cet effet, et validation manuelle. Nous avons focalisé notre analyse sur les adjectifs de forme *inXable* pour plusieurs raisons: a) la préfixation par *in-* est réputée être principalement, voire exclusivement, productive avec des bases en *-able*; b) en même temps, le simple en *-able* fait parfois défaut. Ce dernier point est périodiquement souligné par les linguistes comme étant anormal. Or, notre hypothèse prédit que la situation normale est précisément que soit le simple en *-able*, soit le construit en *in-* fasse défaut, et que, dans le cas contraire, le simple n'exprime pas une propriété à un degré standard. Notre étude sur corpus confirme ces résultats.

Summary

This paper presents the results of a corpus-based analysis of the morphological constructs matching the pattern *in-X-able*, whereby a verbal stem, *X*, is augmented with the suffix *-able*, which derives adjectives from verbs, and with the negative adjectival prefix *in-*. Our objective in this study is twofold: to formulate a theoretical hypothesis regarding the behaviour of the prefixation with *in-* in French; to confront this hypothesis with the constructs attested on a large corpus. In brief, our hypothesis is that in French, *in-* is the exponent of a lexeme construction rule, whose output express the non-satisfaction of the expected property expressed by its input. Our corpus is made of one year of articles extracted from the French newspaper *Le Monde*. We use as well occurrences extracted from articles of other years of *Le Monte* and from the Web. These articles have been first analysed with natural language processing tools; all adjectives matching our pattern of interest being further subject to a manual validation step. The choice of the *in-X-able* constructs was motivated by several reasons: a) prefixation in *in-* is often said to be mainly, if not exclusively, productive with lexemes *X-able*; b) however, in many cases, the hypothesized input of the rule is not observed, if possible at all. This observation has been repeatedly pointed out by linguists as an oddity. Our analysis, on the other hand, predicts that this situation is to be expected, and that the simple lexeme *X-able* and the complex lexeme *in-X-able* are mutually exclusive; when both exist, the *X-able* lexeme does not express an expected property at a standard degree. Our empirical study seems to support these claims.