

L'ORTHOGRAPHE FRANÇAISE EN MOUVEMENT : LE CAS DES EMPRUNTS RECTIFIÉS

RADKA MUDROCHOVÁ, JAN LAZAR

ABSTRACT

French orthography in motion: the case of rectified loanwords

This study examines the integration of rectified spellings of loanwords into contemporary French, with a focus on their presence in lexicographic resources (*Dico en ligne Le Robert* and *Usito*) and their frequency of use in the Aranea web corpora (France and Canada). Based on a selection of 123 loanwords affected by the 1990 French spelling reform, the analysis compares traditional and recommended spellings and observes their distribution across varieties of French. The results confirm a strong representation of English loanwords among the rectified forms, as well as a more systematic implementation of the reform in *Usito* than in *Dico en ligne Le Robert*. Corpus data show that 23% of the lexical items in both French and Canadian corpora occur more frequently in their rectified form, particularly those influenced by English. Although differences between traditional and rectified spellings remain limited overall, the study highlights an ongoing tendency toward stabilization of the reformed orthography. Further research is encouraged to assess the sociolinguistic and normative factors influencing the adoption of rectified spelling in the francophone world.

Keywords: French orthography reform; Loanwords; Lexicography; Corpus linguistics; Language variation; French of France; Canadian French

Mots-clés : Rectifications orthographiques ; emprunts lexicaux ; lexicographie ; linguistique de corpus ; variation du français ; français de France ; français canadien

Introduction

Depuis plusieurs décennies, l'orthographe du français fait l'objet de réflexions et d'ajustements visant à harmoniser le système graphique et à en faciliter l'apprentissage. Parmi ces initiatives, les Rectifications de l'orthographe de 1990 occupent une place particulière, car elles proposent, sans caractère obligatoire, un ensemble de formes recommandées destinées à réduire certaines incohérences graphiques (Lazar, Mudrochová, 2020). Si ces rectifications concernent différents domaines du lexique, elles touchent de manière significative les mots empruntés à d'autres langues, dont l'intégration dans le système ortho-

graphique français peut présenter une série de défis liés à la prononciation, à l'étymologie, à la translittération ou encore au prestige socio-culturel des langues sources.

Or, les emprunts demeurent l'un des moteurs de l'évolution lexicale du français contemporain, notamment dans des champs tels que la technologie, la culture médiatique, l'alimentation ou les pratiques sociales. De ce fait, l'étude des formes rectifiées constitue un observatoire important pour comprendre comment une langue adapte et transforme les lexies venues d'ailleurs afin de les intégrer à son propre fonctionnement grapho-phonologique.

La présente recherche s'appuie sur un corpus de 123 emprunts identifiés à partir de la base du *RENOUVO* (Réseau pour la nouvelle orthographe du français) et examinés dans plusieurs ressources complémentaires : les dictionnaires *Dico en ligne Le Robert*¹ et *Usito*, ainsi que les corpus web *Aranea* (France et Canada). Elle vise à déterminer dans quelle mesure les graphies recommandées circulent dans l'usage réel et comment elles sont prises en charge par les principales sources lexicographiques de référence. L'analyse accorde également une attention particulière aux variétés du français, afin d'observer d'éventuels écarts entre usage de France et usage canadien, notamment dans le traitement des anglicismes.

L'objectif est ainsi de contribuer à une meilleure compréhension des enjeux liés à l'évolution de l'orthographe française et, plus précisément, au statut des emprunts dans ce contexte de changement orthographique progressif.

1. Bref aperçu historique

L'histoire de l'orthographe française est un récit fascinant d'évolutions, de querelles et de tentatives de standardisation, oscillant constamment entre une approche phonétique, l'influence latinisante et, finalement, la recherche d'une norme fixée par des institutions. Avant de passer à la présentation des emprunts rectifiés, nous jugeons utile de présenter un bref aperçu historique de son évolution.

Il faut se rendre compte qu'au début, l'orthographe de la langue française était rudimentaire et non standardisée. Le français, issu du latin vulgaire, est apparu pour la première fois à l'écrit dans les Serments de Strasbourg en 842. À cette époque-là, l'écrit était dominé par le latin, la langue de l'Église et de l'administration. Soulignons qu'au Moyen Âge (jusqu'au XIII^e siècle), l'orthographe de la langue française était principalement phonétique. Les copistes et les écrivains notaient les mots tout simplement comme ils les prononçaient. Cette absence de norme entraînait une grande variabilité : un même mot pouvait être écrit de plusieurs façons différentes dans le même document, voire par la même personne (Pellat, 2023 : 10). Il en résulte qu'à l'époque médiévale, il n'existait pas

¹ Le *Dico en ligne Le Robert* (version gratuite 2025) a été privilégié par rapport à sa version complète payante pour plusieurs raisons. D'une part, ce choix permet d'assurer une cohérence méthodologique avec le dictionnaire *Usito*, également accessible librement en ligne. D'autre part, il garantit la reproductibilité de l'analyse, les ressources mobilisées étant directement consultables par tout lecteur sans restriction d'accès. Enfin, l'étude s'intéresse également au traitement des graphies dans des interfaces lexicographiques destinées à un large public ; à ce titre, le recours à des versions librement accessibles reflète plus fidèlement les conditions réelles de consultation des usagers, dont les pratiques constituent un élément pertinent pour l'analyse.

une seule norme orthographique, mais le même mot disposait de plusieurs variantes. L'objectif n'était pas de fixer les règles abstraites orthographiques, mais la compréhension immédiate par la lecture. La période du moyen français (XIII^e–XV^e siècles) a vu un changement majeur (Pellat, 2023 : 34). L'introduction du français dans les actes publics et le développement de l'imprimerie ont multiplié les documents écrits, ce qui a créé la nécessité de régularisation de l'orthographe (Helix, 2011 : 250). Cette période se caractérisait par une grande instabilité et par des débats passionnés autour de l'orthographe. Les deux tendances principales s'affrontaient :

- a) Les phonétistes (comme Louis Meigret) prônaient une orthographe fidèle à la prononciation.
- b) Les étymologistes (ou « latinisants ») voulaient « ennoblir » la langue française en y réintroduisant des lettres muettes, souvent inspirées de l'orthographe latine d'origine des mots, pour des raisons de prestige culturel. Par exemple, le mot *doigt* (du latin *digitus*) s'est vu ajouter un *g* et un *t* muets. De mêmes exemples fournissent les mots *temps* ou *corps* qui contiennent un *p* et *s* muets.

Au XVII^e siècle s'est mise en place une standardisation de la langue française. La création de l'Académie française en 1635 a joué un rôle crucial dans ce processus de la standardisation. Le premier *Dictionnaire de l'Académie française*, publié en 1694, vise à établir une norme prescriptive. Malheureusement, les Académiciens ont opéré des choix, souvent conservateurs, figeant des graphies qui ne correspondaient déjà plus à la prononciation de l'époque (Helix, 2011 : 263). Il est possible de constater que ces tendances conservatrices se sont poursuivies jusqu'au XIX^e siècle. Pourtant, c'est l'instruction obligatoire laïque et gratuite (instituée par la loi du 28 mars 1882, dite « loi Jules Ferry ») qui a changé la situation. On s'est rapidement rendu compte que l'orthographe non-transparente constituait un vrai défi pour la population et une grande partie des élèves n'était pas en mesure de maîtriser pleinement les subtilités de ce système graphique si exigeant. La « faute d'orthographe » est ainsi devenue, plus que jamais, un marqueur social d'éducation et de prestige. On observait aussi les premières tentatives de simplifier l'orthographe pour le rendre plus accessible à tout le monde. Les modifications de l'orthographe du XIX^e siècle visaient principalement à clarifier et à harmoniser les règles établies, notamment en matière d'accents et de traits d'union (Helix, 2011 : 266). À titre d'exemples, nous pouvons mentionner les mots comme *poète* > *poète* (le tréma est remplacé par un accent grave) ou *collège* > *collège* (le *é* est remplacé par *è* pour mieux correspondre à la prononciation). Le XX^e siècle a été marqué par des débats récurrents sur la simplification de l'orthographe. Les rectifications de 1990, proposées par le Conseil supérieur de la langue française et approuvées par l'Académie française, sont l'exemple le plus notable. Vu que l'objectif de notre article est d'analyser les graphies rectifiées des mots empruntés, nous désirons les présenter plus en détail dans les paragraphes suivants.

2. Rectifications orthographiques (le cas de l'emprunt)

Les rectifications orthographiques de 1990 ont apporté des modifications significatives visant à rationaliser et simplifier l'orthographe française. Une partie importante de ces changements concerne les mots empruntés. L'objectif principal est de les franciser, c'est-

à-dire de les adapter aux règles orthographiques et grammaticales du français, pour les intégrer plus harmonieusement dans la langue.

Il est possible de distinguer plusieurs axes :

- a) Application des règles d'accentuation françaises – de nombreux mots d'origine étrangère, en particulier ceux qui contiennent un *e* prononcé [e] ou [ɛ], n'avaient pas d'accent dans leur graphie d'origine. Les rectifications ont introduit l'utilisation des accents aigus et graves pour se conformer à la norme française et permettre une meilleure indication de la prononciation correcte en français.

sombrero → *sombréro*

tequila → *téquila*

allegro → *allégro*

pietà → *piéta*

- b) Régularisation du pluriel – il est recommandé de former le pluriel selon la règle française (ajout d'un *s* ou *x*), ce qui met fin aux pluriels étrangers complexes.

scenarii → *scénarios*

maffiosi → *des mafiosos*

des maxima → *des maximums*

gentlemen → *des gentlemen*

matches → *des matchs*

- c) Soudure – il est recommandé de souder les mots composés courants.

baby-foot → *babyfoot*

hand-ball → *handball*

tee-shirt ou *T-shirt* → *teeshirt*

- d) Francisation de la graphie – cette catégorie regroupe les emprunts dont l'orthographe a été modifiée pour mieux correspondre aux règles phonétiques et graphiques du français.

électropuncture → *électroponcture*

kirsh → *kirch*

uléma → *ouléma*

letchi → *litchi*

Afin d'examiner concrètement l'impact de ces rectifications sur les emprunts lexicaux, il est nécessaire de passer d'une présentation théorique à une démarche méthodologique permettant d'analyser systématiquement les formes concernées. La section suivante décrit donc les étapes de constitution du corpus et les critères de sélection appliqués pour identifier, vérifier et classer les emprunts rectifiés. Elle permet ainsi de traduire les orientations générales des rectifications de 1990 en un dispositif d'analyse rigoureux, fondé sur des ressources lexicographiques et des critères linguistiques clairement définis.

3. Méthodologie de recherche

Le point de départ de cette étude est la liste des mots présentés sur le site du *RENOUVO* (Réseau pour la nouvelle orthographe du français), qui rassemble les formes recommandées dans le cadre de la nouvelle orthographe. Cette ressource répertorie plus de 2 500 mots concernés par les recommandations orthographiques.

À partir de cette base, une première sélection ciblée a été réalisée en fonction de la règle G2, qui porte sur l'adaptation et la francisation des emprunts lexicaux. L'application de cette règle a permis de constituer un corpus de 123 unités lexicales. Selon cette règle, les rectifications orthographiques recommandent de manière générale aux auteurs et autrices de dictionnaires, ainsi qu'aux créateurs et créatrices de mots, de franciser, dans la mesure du possible, les emprunts, en les adaptant au système graphique du français (par exemple : *débatteur*, *musli*). Par ailleurs, lorsque plusieurs formes coexistent, il est conseillé d'accorder la préférence à celle qui est la plus conforme au fonctionnement grapho-phonologique du français (par exemple : *paélia*, *taliatelle*).

Dans un second temps, chacune des lexies sélectionnées a fait l'objet d'une vérification systématique dans plusieurs ressources lexicographiques de référence : *Dico en ligne Le Robert* (version gratuite 2025, disponible en ligne) et *Usito*, auxquels s'ajoute le *Dictionnaire de l'Académie française*, particulièrement mobilisé pour l'origine étymologique des emprunts et, le cas échéant, pour les indications relatives aux formes concurrentes.

Les unités lexicales ont ensuite été réparties selon deux critères principaux :

1. la langue source de l'emprunt, permettant d'identifier les aires linguistiques les plus productives (anglais, italien, espagnol, allemand, *etc.*) ;
2. la partie du discours (nom, verbe, adjectif, adverbe), afin d'observer d'éventuelles tendances liées à la catégorie grammaticale.

Cette organisation a permis de commenter les différents procédés de francisation en fonction des langues d'origine, en mettant en évidence les régularités et les cas particuliers.

Enfin, pour mesurer la présence effective et la diffusion des formes étudiées, les occurrences ont été vérifiées dans les corpus web *Aranea* (pour en savoir plus, *cf.* notamment Benko 2018 ; Benko, Butašová et Puchovská 2019), développés par l'Académie slovaque des sciences (plateforme : <http://unesco.uniba.sk>). Ils regroupent principalement des textes contemporains issus de divers genres discursifs (presse en ligne, blogs, forums, pages institutionnelles, contenus informatifs et promotionnels), reflétant ainsi une large diversité d'usages. Les données sont collectées par campagnes de téléchargement sur le web, généralement sur une période temporelle limitée afin d'assurer la comparabilité des corpus ; par exemple, le corpus *Araneum Francogallicum* repose sur des données collectées en 2013, bien que la couverture temporelle puisse varier selon les versions et mises à jour des différents sous-corpus. L'utilisation parallèle de deux grands corpus comparables, *Araneum Francogallicum Gallicum* (≈ 3,29 milliards de mots) et *Araneum Francogallicum Canadiense II* (≈ 406 millions de mots), a permis de distinguer les usages selon les variétés du français, en particulier entre le contexte hexagonal et le contexte canadien.

Dans un cas particulier, à savoir celui de la lexie *cashier*, qui présente un nombre particulièrement élevé de variantes graphiques, l'analyse a été complétée par le recours à deux outils permettant d'introduire une dimension diachronique. D'une part, la plateforme *Google Ngram Viewer*, qui permet de suivre l'évolution de la fréquence des formes dans de vastes corpus de livres numérisés, a été utilisée. D'autre part, la plateforme *Nénufar* (*cf.* Bohbot, Frontini, Faucher, 2019), qui donne accès à des éditions anciennes du *Petit Larousse illustré*, a permis de repérer des attestations lexicographiques et de documenter l'ancienneté de certaines graphies.

4. Présentation des résultats

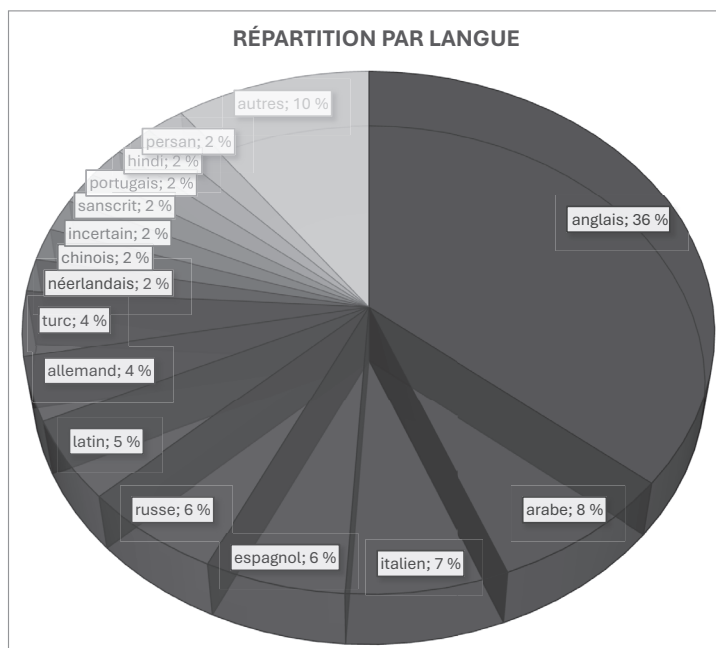
Dans ce chapitre, nous cherchons à décrire les propriétés morphosyntaxiques, étymologiques et graphiques des lexies étudiées, ainsi que leur traitement lexicographique et leur fréquence en usage. Une attention particulière est accordée à la variation orthographique liée aux rectifications, envisagée en relation avec le degré d'intégration des emprunts en français.

4.1 Questions morphosyntaxique et étymologique

L'analyse morphosyntaxique du corpus met en évidence une dominance très nette des noms, qui représentent 106 occurrences. On relève également 7 unités pouvant fonctionner à la fois comme nom et adjectif, ainsi que 3 adjectifs. Le corpus comprend en outre 4 verbes (*iodler, shampooiner, snifer, chlinguer*), 2 interjections (*snif, tcha(o)*) et 1 locution (*à capella*).

Cette répartition confirme que les emprunts francisés concernent majoritairement des noms, ce qui s'explique par la forte proportion de désignations d'objets, d'activités, de réalités culturelles ou gastronomiques introduites par contact linguistique. Les verbes et les interjections restent plus rares, mais témoignent de phénomènes d'appropriation expressive et pragmatique.

Pour ce qui est des langues d'origine des emprunts, leur répartition est présentée par le graphique 1.



Graphique 1 Répartition des langues d'origine des emprunts étudiés

Le graphique montre que l'anglais constitue la source la plus importante, représentant 36 % des lexies recensées. Viennent ensuite les emprunts à l'arabe (8 %) et à l'italien (7 %), suivis de l'espagnol et du russe, chacun comptant pour 6 % du corpus. Les apports issus du latin correspondent à 5 %, tandis que les emprunts à l'allemand et au turc représentent chacun 4 %.

Les langues ayant une contribution plus marginale sont le néerlandais, le persan, le portugais, le hindi, le sanskrit et le chinois, ainsi que certaines étymologies incertaines, chacune de ces catégories étant estimée à environ 2 %. Enfin, la catégorie « autres » totalise 10 %, regroupant un ensemble de langues dont la présence reste plus marginale mais témoigne de la richesse et de la diversité des sources d'emprunts. On y trouve notamment des lexies provenant de l'hébreu, du hongrois, de l'inuit, du japonais, du nahuatl, du norvégien, du quechua, de langues dites scandinaves, ensuite du soudanais, du suédois, de l'alémanique, du tibétain et du tartare.

Dans l'ensemble, cette répartition confirme le poids déterminant de l'anglais dans les emprunts intégrés au français contemporain, tout en montrant une variété importante d'apports lexicaux provenant de différentes aires linguistiques.

4.2 Questions lexicographiques

L'examen du corpus dans *Dico en ligne Le Robert* révèle une forte hétérogénéité dans le traitement lexicographique des emprunts soumis aux rectifications orthographiques. Deux lexies n'y apparaissent pas du tout, à savoir *séniorita* et *dermoponcture*. Dans quatre autres cas, seule la graphie non rectifiée est attestée, comme pour *yiddish*, *Ostrogoth*, *téocalli* ou encore *llanos*. À l'inverse, six entrées ne sont représentées que sous leur graphie rectifiée, notamment *calo*, *mamelouk*, *débatteur*, *ouzbek*, *gourou* et *mixeur*. Pour le reste du corpus, les deux formes, voire davantage, sont attestées. C'est le cas, par exemple, de *malstrom*, pour lequel le dictionnaire recense jusqu'à quatre variantes concurrentes, dont *maelstrom*, *malström* et *maelström*. La présentation de ces variantes n'est toutefois pas homogène. Dans de nombreuses notices, *Dico en ligne Le Robert* ne signale aucune préférence entre les formes traditionnelles et celles proposées par les rectifications de l'orthographe. L'article consacré à *globe-trotteur* (cf. la figure 1) en offre une illustration évidente, puisqu'il réunit plusieurs graphies, dont la forme rectifiée *globetrotteur*.

En revanche, l'examen du corpus dans le dictionnaire *Usito* montre une présence de l'orthographe rectifiée plus systématique que dans *Dico en ligne Le Robert*. On relève

globe-trotteur , globe-trotteuse nom

vieilli Voyageur qui parcourt la terre. Des globe-trotteurs.

• **var.** globe-trotter | nom masculin **RO** globetrotteur | globetrotteuse

Figure 1. Présentation de la lexie *globetrotteur* dans le PR

cependant dix-sept lexies complètement absentes du dictionnaire, parmi lesquelles p. ex. *bogie*, *dealeur*, *boscop* ou encore *dinghie*. Un seul mot n'y figure qu'avec sa graphie non rectifiée, soit *muezzin*, tandis que cinq entrées ne sont attestées qu'avec la graphie issue de la réforme, notamment *gourou*, *halal*, *talweg*, *ouzbek* et *mamelouk*. Dans certains cas, *Usito* propose deux graphies considérées comme recommandées, par exemple *cougar* ou *cougouar*, choix qui dépend de la prononciation. Le traitement des anglicismes² se révèle plus spécifique encore, les deux graphies étant souvent indiquées avec une remarque sur l'emploi critiqué de l'emprunt. L'article *shaker* en fournit une illustration : en plus de la graphie rectifiée *shakeur*, le dictionnaire précise que l'usage du mot *shaker* (cf. la figure 2) est parfois jugé non standard en contexte francophone, et que le terme *coquetelier*, présenté comme équivalent, demeure peu attesté dans l'usage. Il convient toutefois de souligner que cette remarque ne relève pas des rectifications orthographiques à proprement parler, mais concerne l'évaluation de l'emprunt lui-même, indépendamment de sa graphie ; cette distinction est particulièrement importante dans le contexte québécois, où les anglicismes font l'objet d'une attention normative spécifique et de recommandations explicites visant à encadrer leur usage, comme évoqué précédemment.

shaker [ʃɛkœʁ] n. m.

| RO *shakeur*.

REM. L'emploi de *shaker* est parfois critiqué comme synonyme non standard de *coquetelier*. Cet équivalent est peu attesté dans l'usage.

Figure 2 Présentation de la lexie *shaker* dans le dictionnaire *Usito*

4.3 Recherche de fréquence – corpus *Aranea*

La consultation du corpus *Araneum* (d'abord de France) montre que, de manière générale, l'orthographe traditionnelle demeure majoritaire, à l'exception de 28 lexies (présentées dans le tableau 1) pour lesquelles la forme issue des rectifications de l'orthographe s'avère plus fréquente, le nombre d'occurrences (la fréquence absolue) apparaît après chaque forme.

² Ce fonctionnement correspond aux principes descriptifs et normatifs du dictionnaire *Usito*, cf. le chapitre destiné uniquement aux « anglicismes critiqués » : https://usito.usherbrooke.ca/index/anglicismes_critiqués, où les emprunts à l'anglais font l'objet de commentaires explicites sur leur acceptabilité et leur usage, ainsi que de suggestions de formes recommandées. Ce type de traitement s'inscrit plus largement dans une tradition lexicographique québécoise attentive à la variation et à la norme, telle qu'analysée notamment par Nadine Vincent (cf. Sorba, Vincent 2025 ; Remysen, Vincent, 2017) qui souligne le rôle des dictionnaires dans la médiation entre usage réel et recommandations linguistiques, en particulier dans le domaine des anglicismes.

Tableau 1 Lexies à graphie recommandée prédominante dans *Araneum* (France)

RO FR	Traditionnelle FR	RO FR	Traditionnelle FR
un aga 191	un agha 5	mafioso 210	maffioso 14
angstroem 14	angström 0	mamelouk 892	mameluk 132
aquavit 597	akvavit 13	marijuana 2582	marihuana 96
bogie 1089	boggie 213	mixeur 1177	mixer 669
calo 208	caló 18	moukère 55	mouquère 11
canyon 12347	cañon 100	oukase 644	ukase 192
cashier 1705	Kasher 259, cascher 16, cachère 250, cascher 16, kasher 259, kascher 70	ouléma 882	uléma 289
débatteur 557	debater 33	ouzbek 1238	uzbek 56
fioul 13 305	fuel 6649	shampoing 18717	shampooing 10893
gourou 8742	guru 1581	shampouiner 157	shampooiner 40
halal 5673	hallal 1852	shampouineur 147	shampooineur 9
hippie 4731	hippy 293	snif 3200	sniff 1095
litchi 1679	letchi 260	talweg 562	thalweg 386
mafia 11380	maffia 520	yak 1162	yack 920

Les anglicismes occupent une place notable dans cet ensemble : ils représentent 32 % des emprunts relevés dans le corpus français, proportion relativement proche des 36 % observés dans le corpus total. Certains écarts sont particulièrement significatifs lorsque l'on compare les données du français hexagonal avec celles du corpus canadien. Il convient toutefois de préciser que ces différences ne concernent pas uniquement des variantes graphiques, mais relèvent également de choix lexicaux distincts selon les aires géographiques. Ainsi, la lexie *fioul*, nettement plus fréquente dans l'usage français de France, correspond à l'emploi privilégié de *mazout* au Québec. Le tableau met en évidence de nombreux cas où la graphie rectifiée circule plus largement que la graphie traditionnelle. Ainsi, on relève des occurrences nettement plus élevées pour des formes comme *fioul* (13 305 occurrences contre 6 649 pour *fuel*), *gourou* (8 742 contre 1 581 pour *guru*), *halal* (5 673 contre 1 852 pour *hallal*), *ouzbek* (1 238 contre 56 pour *uzbek*), ou encore *mafioso* (210 contre 14 pour *maffioso*), par rapport à leurs variantes dites plus traditionnelles.

La recherche dans du le corpus *Araneum* canadien, fondé sur des sources web en français canadien, montre que l'orthographe traditionnelle demeure dominante, à l'exception de 28 lexies pour lesquelles la graphie issue des rectifications orthographiques apparaît plus fréquente (cf. le tableau 2).

Les anglicismes y occupent une position encore plus marquée que dans le corpus de France. Ils représentent en effet 39 % des emprunts analysés, proportion légèrement supérieure à aux 36 % observés dans l'ensemble du corpus. Les données témoignent également de différences graphiques sensibles avec le français de France, notamment en ce qui concerne les emprunts à l'anglais, comme en attestent *cornflakes* (5 occurrences

Tableau 2 Lexies à graphie recommandée prédominante dans *Araneum* (Canada)

RO CA	Traditionnelle CA	RO CA2	Traditionnelle CA2
cornflakes 5	corn-flakes 3	mafia 721	maffia 10
pouding 604	pudding 139	mafioso 28	maffioso 2
yogourt 3752	yaourt 389, yoghourt 25	mamelouk 27	mameluk 6
un aga 16	un agha 0	marijuana 1971	marihuana 367
angstrøm 0	angström 0	mixeur 25	un mixer 3
aquavit 4	akvavit 2	moukère 1	mouquère 0
bogie 200	boggie 7	oukase 8	ukase 7
calo 22	caló 1	ouléma 55	uléma 17
canyon 982	cañon 4	ouzbek 330	uzbek 3
casher 312	kasher 60, cascher 4, cachère 136, kascher 27	shampoing 466	shampooing 431
débatteur 79	debater 17	shampouiner 5	shampooiner 0
gourou 614	guru 159	shampouineur 9	shampooineur 0
halal 946	hallal 46	snif 98	sniff 19
hippie 296	hippie 20	yak 45	yack 29
litchi 100	letchi 1		

contre 3 pour *corn-flakes*), *pouding* (604 occurrences, contre 139 pour *pudding*) ou encore *yogourt* (3 752 occurrences, face à 389 pour *yaourt* et 25 pour *yoghourt*). En outre, plusieurs cas confirment toutefois la préférence canadienne pour certaines graphies rectifiées largement répandues, comme *canyon* (982 occurrences), *casher* (312), *débatteur* (79), *gourou* (614), *halal* (946), *hippie* (296) ou *litchi* (100), formes qui dépassent nettement leurs équivalents traditionnels.

Les résultats montrent que si l'orthographe rectifiée tend à s'imposer pour certains emprunts, l'orthographe traditionnelle demeure dominante pour d'autres. Cette variation semble s'expliquer par plusieurs facteurs, notamment le degré de diffusion des formes dans l'usage, leur fréquence, ainsi que les pratiques discursives dans lesquelles elles s'inscrivent. Elle peut également être influencée par des facteurs normatifs ou diatopiques, ainsi que par la visibilité socioculturelle de certaines formes, qui contribuent à maintenir une concurrence graphique durable.

Il convient toutefois de souligner que ces interprétations reposent exclusivement sur des données issues de corpus écrits. Une analyse plus approfondie, notamment fondée sur des enquêtes auprès de locuteurs natifs, serait nécessaire pour mieux cerner les mécanismes en jeu et évaluer la perception et l'acceptabilité des différentes graphies. Dans cette étude, nous nous limitons ainsi à une approche descriptive visant à cartographier les usages observés dans les données disponibles.

4.4 Cas particulier – *casher*

Pour terminer cette partie de notre analyse et illustrer plus concrètement la variabilité graphique observée dans l'adaptation des emprunts, il est utile de s'arrêter sur un cas particulièrement significatif. Le mot *casher* en constitue en effet un exemple représentatif. Il s'agit de la lexie qui présente le plus grand nombre de variantes orthographiques attestées dans les corpus consultés, parmi lesquelles *kasher*, *cascher*, *cachère*, *cascher*, *kasher*, *kascher* ou encore *cawcher* (1925) ou *kawcher* (1948), les deux derniers exemples relevés sur la plateforme *Nénufar*.

L'exploration diachronique de ces formes au moyen de *Google Ngrams* (cf. la figure 3) confirme la prédominance très nette de la graphie *casher* depuis la fin du XX^e siècle, alors que *kasher* et *cachère* se maintiennent à des niveaux beaucoup plus faibles.

Cette évolution révèle un processus de stabilisation orthographique allant dans le sens des recommandations de la rectification, tout en montrant que la concurrence graphique demeure historiquement ancrée et encore perceptible dans certains usages spécialisés ou communautaires.

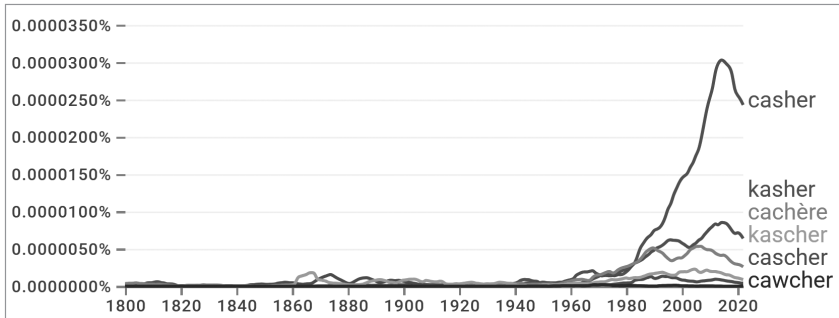


Figure 3 Fréquence diachronique des variantes orthographiques de *casher* dans Google Ngram Viewer

Conclusion

L'ensemble des observations menées dans cette étude montre que les emprunts à l'anglais demeurent les plus représentés parmi les lexies concernées par les rectifications orthographiques, ce qui s'inscrit dans une tendance bien établie, l'anglais constituant depuis plusieurs décennies la principale langue source d'emprunts en français.

Sur le plan lexicographique, la présentation des graphies réformées apparaît plus systématique dans le dictionnaire *Usito*, tandis que *Dico en ligne Le Robert* intègre un nombre plus important d'emprunts issus de notre corpus, sans que la relation entre les variantes soit toujours explicitement précisée. Certaines lexies révèlent des divergences normatives entre les ressources consultées.

Du côté des usages en ligne, 23 % des lexies du corpus français présentent une fréquence plus élevée pour la graphie rectifiée, proportion identique à celle observée dans le corpus canadien, où la présence plus marquée d'anglicismes contribue à renforcer certaines formes rectifiées.

Si les différences d'usage entre graphies traditionnelles et graphies rectifiées restent limitées dans les deux corpus, les données montrent néanmoins que certaines formes recommandées présentent des fréquences plus élevées, ce qui indique une diffusion non négligeable dans l'usage contemporain.

Ces résultats invitent à poursuivre la recherche, notamment en élargissant le corpus à d'autres variétés du français, en affinant les mesures de fréquence (par exemple via une normalisation statistique) et en intégrant d'autres sources lexicographiques ou d'usage. Une analyse complémentaire, attentive aux dimensions sociolinguistiques, terminologiques et éditoriales, permettra de mieux comprendre les conditions de diffusion et d'acceptabilité des rectifications orthographiques dans l'usage du français actuel.

BIBLIOGRAPHIE

- Beaulieux, Charles (1927–1967). *Histoire de l'orthographe française*. Paris : Champion.
- Benko, Vladimír (2018). « Aranea : A Family of Comparable Gigaword Web Corpora ». Dans *Web Corpora & Corpus Linguistics Portal*. Bratislava : Slovak Academy of Sciences L. Štúr Institute of Linguistics, http://aranea.juls.savba.sk/aranea_about/index.html [consulté le 21/01/2025].
- Benko, Vladimír – Butašová, Anna – Puchovská, Zuzana (éds.) (2019). *Webové korpusy Aranea. Učebnica pre učiteľov jazykov, prekladateľov, tlmočníkov, filológov a študentov filologických odborov*. Bratislava : Univerzita Komenského.
- Bohbot, Hervé – Frontini, Francesca – Faucher, Alexandre (2019). « Le projet Nénufar, nouvelle édition numérique du Petit Larousse illustré (1906–1948) ». Dans *XXIX^e congrès international de linguistique et de philologie romanes*. Copenhague : Société de Linguistique Romane. <https://hal.science/hal-03262248> [consulté le 21/01/2025].
- Catach, Nina (1973). « La structure de l'orthographe française », *La Recherche*, 39(4), p. 949–956.
- Catach, Nina (2001). *Histoire de l'orthographe française*. Paris : Champion.
- Contant, Chantal – Muller, Romain (2010). *Les rectifications de l'orthographe du français*. Bruxelles : De Boeck / Éditions Duculot.
- Conseil supérieur de la langue française (1990). *Les Rectifications de l'orthographe*. *Journal officiel de la République française*, Documents administratifs, n° 100, 6 décembre.
- Chervel, André (2008). *L'orthographe en crise à l'école : Et si l'histoire montrait le chemin ?* Paris : Retz.
- Dister, Anne – Moreau, Marie-Luce (2012). « Réforme de l'orthographe française. Craintes, attentes et réactions des citoyens – Présentation », *Glottopol*, n° 12, en ligne http://glottopol.univ-rouen.fr/numero_19.html [consulté le 21/01/2025].
- Goosse, André (1991). *La « nouvelle » orthographe. Exposé et commentaires*. Paris / Louvain-La-Neuve : Duculot.
- Hélix, Laurence (2011). *Histoire de la langue française*. Paris : Ellipses.
- Jacquet-Pfau, Christine (2011). « Les emprunts lexicaux dans la neuvième édition du Dictionnaire de l'Académie française », *ÉLA (Études de linguistique appliquée)*, n° 163, p. 307–323.
- Lazar, Jan – Mudrochová, Radka (2020). « Les rectifications orthographiques de 1990 et leur application dans l'espace francophone, 30 ans plus tard », *Cahiers de Praxématique*, n° 74, en ligne <https://journals.openedition.org/praxematique/6098> [consulté le 21/01/2025].
- Pellat, Jean-Christophe (2023). *L'orthographe française : histoire, description, enseignement*. Paris : Orphys.
- Remysen, Wim – Vincent, Nadine (dir.) (2017). La lexicographie au Québec. *Revue de l'Université de Moncton*, 48(1).
- Sorba, Nicolas – Vincent, Nadine (dir.) (2025). Les dictionnaires. *Ponts. Langues, littératures, civilisations des pays francophones*, 25.
- Walter, Henriette (1998). *L'Aventure des mots venus d'ailleurs*. Paris : Robert Laffont.

SITOGRAPHIE

- Aranea Web Corpora. *Aranea – Web Corpora & Corpus Linguistics Portal*, <http://unesco.uniba.sk> [consulté le 21/01/2025].
- Dictionnaire de l'Académie française. *Neuvième édition du Dictionnaire de l'Académie française* [dictionnaire en ligne], <https://www.dictionnaire-academie.fr> [consulté le 23/03/2025].
- Google Books Ngram Viewer. *Google Ngram Viewer*, <https://books.google.com/ngrams> [consulté le 21/01/2023].
- Le Robert. *Dico en ligne Le Robert*, <https://dictionnaire.lerobert.com> [consulté le 22/09/2025].
- Nénufar. *Nénufar – Nouvelle édition numérique de fac-similés de référence (Petit Larousse illustré 1905–1948)*, <https://nenufar.huma-num.fr> [consulté le 21/01/2023].
- RENOUVO. *Réseau pour la nouvelle orthographe du français*, <http://www.renouvo.org> [consulté le 09/01/2025].
- TLFi. *Trésor de la langue française informatisé (TLFi)* [portail CNRTL], <https://atilf.atilf.fr> [consulté le 21/01/2023].
- Usito. *Usito, dictionnaire de la langue française*, <https://usito.usherbrooke.ca> [consulté le 22/07/2025].

Radka Mudrochová
Institut d'Études Romanes
Université Charles
nám. Jana Palacha 2, 116 38, Prague 1, Tchéquie
radka.mudrochova@ff.cuni.cz
ORCID : 0000-0002-8718-6922

Jan Lazar
Département d'Études Romanes
Université d'Ostrava – Université d'Opole
Reální 5, Ostrava, Tchéquie – Plac Kopernika 11a, Opole, Pologne
NEOLAiA / Le Studium, Loire Valley Institute for Advanced Studies, Orléans
& Tours, France – IRJI François-Rabelais EA 7496 – Institut de Recherche Juridique
Interdisciplinaire François-Rabelais (University of Tours), France
janlazar@email.cz
ORCID : 0000-0002-2436-7152