



HAL
open science

Quantité et qualité dans le Wiktionnaire : de la diversité... à la rigueur ?

Franck Sajous

► **To cite this version:**

Franck Sajous. Quantité et qualité dans le Wiktionnaire : de la diversité... à la rigueur ?. Linx, 2023, 86, 10.4000/linx.9835 . halshs-04273590v2

HAL Id: halshs-04273590

<https://shs.hal.science/halshs-04273590v2>

Submitted on 24 Nov 2023

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.

Quantité et qualité dans le *Wiktionnaire* : de la diversité... à la rigueur ?

Franck Sajous
CLLE, CNRS & Université de Toulouse 2

Résumé : Cet article porte sur le *Wiktionnaire* et s'intéresse plus précisément aux liens entre (manque de) révisions et (manque de) qualité. Ces liens sont étudiés d'abord d'un point de vue quantitatif, puis qualitatif. L'analyse quantitative du taux de révision des articles, et plus particulièrement des définitions, dément la croyance répandue que le dictionnaire contributif est abondamment révisé par une grande diversité de contributeurs. Il apparaît au contraire que les révisions sont le fait d'une poignée de contributeurs et se concentrent sur un faible nombre de pages. L'article montre également que les révisions – quand elles existent – concernent principalement l'ajout de contenu ou des modifications de surface, plus rarement la modification du texte existant, et encore moins la modification des définitions. La seconde analyse, qualitative, s'intéresse à une variété de problèmes présents dans les descriptions lexicographiques d'entrées de différentes natures (entrées polémiques, termes spécialisés ou informations factuelles) et, pour chaque cas, étudie le nombre et la nature des révisions. Cette analyse montre d'une part que la persistance de certaines erreurs sont dues à un manque de révision, tout en contredisant, d'autre part, le présupposé partagé par plusieurs travaux antérieurs postulant que plus de révisions impliquent une meilleure qualité (en règle générale, elles produisent plus de contenu, *i.e.* plus de quantité, mais pas une meilleure qualité). De ces observations découle un constat sur le processus rédactionnel du *Wiktionnaire* permettant d'affirmer que ce dictionnaire contributif est plus agrégatif que collaboratif.

Mots-clés : *Wiktionnaire*, édition collaborative, révision des définitions, rigueur, diversité

Pour citer cet article :

Franck Sajous. (2023). Quantité et qualité dans le *Wiktionnaire* : de la diversité... à la rigueur ? *Linx*, 86.

COPYRIGHT

Ce document est la version auteur de l'article. Il est distribué sous licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International (CC BY-NC-ND 4.0)



<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.fr>

La version de la revue *Linx* est disponible en ligne : <https://doi.org/10.4000/linx.9835>

Quantité et qualité dans le *Wiktionnaire* : de la diversité... à la rigueur ?

Franck Sajous
CLLE, CNRS & Université de Toulouse 2

*“In this democratic days, any investigation
into the trustworthiness and peculiarities of
popular judgments is of interest”*

Francis Galton, *Vox Populi*

*« il y avait cette sacro-sainte et despotique
opinion publique, c'est-à-dire l'opinion de
l'Homme Moyen, [...] laquelle (opinion)
favorisait un extraordinaire bouillonnement
de médiocrités »*

Lydie Salvayre, *Hymne*

1 Introduction

Lors de la seconde édition du colloque *Dictionnaires et culture numérique dans l'espace francophone*, Sajous et al. (2018a) s'intéressaient au principe de neutralité de point de vue revendiqué par le *Wiktionnaire*¹ et montraient que ce dictionnaire n'est pas toujours fidèle à la ligne de conduite qu'il se fixe. Les écarts relevés dans l'étude se manifestent notamment par l'utilisation d'axiologiques subjectifs dans les définitions de noms de vins et de cépages et par l'alternance entre marques de distance et de prise en charge dans le domaine de l'occultisme. Pour les articles analysés, les entorses au principe de neutralité relevées sont essentiellement imputables à l'import automatique d'articles de dictionnaires anciens, tels que le *Littré* et la 8^e édition du *Dictionnaire de l'Académie française* (DAF8), et surtout à une absence de révision des définitions par les contributeurs du *Wiktionnaire*. Le jour même du colloque, un wiktionnariste présent dans l'assistance supprimait les évaluatifs des trois définitions de vins citées. Un an plus tard, lors de la publication de l'article faisant suite au colloque, ce même wiktionnariste récidivait en modifiant les définitions constituant le jeu de données fourni par Sajous et al. (2019): suppression d'évaluatifs

1. *Wiktionnaire* désigne l'édition française du dictionnaire contributif. *Wiktionary* désigne à la fois sa version anglaise et l'ensemble du projet (*i.e.* toutes les éditions de langue réunies). Les articles de chaque édition sont susceptibles d'être modifiés à tout moment. Dans la présente étude, les études quantitatives de la section 5 sont fondées sur une version des articles au 1/09/2021. Sauf mention contraire, les descriptions qualitatives ont été menées au printemps 2022. L'historique des versions d'un article donné est accessible depuis le lien « Voir l'historique » en haut à droite de la page correspondante.

dans le domaine du vin et introduction systématique de marques de distance dans celui de l'occultisme. Ces choix sont compréhensibles car ils visent à mettre le *Wiktionnaire* en conformité avec un de ses principes fondateurs. Ils sont néanmoins discutables à plusieurs titres. D'une part, en supprimant certains jugements subjectifs des définitions de vins et de cépages, le contributeur supprimait également des éléments informatifs. D'autre part, concernant le domaine de l'occultisme, une volonté d'homogénéisation aurait pu conduire aussi bien à une suppression qu'à un ajout systématique des marques de distance. Enfin, pour les deux domaines, il reste dans le *Wiktionnaire* des entrées, non mentionnées par Sajous et al. (2019), dans lesquelles on trouve encore axiologiques subjectifs et prises en charge. Au-delà de l'anecdote, cette séquence est symptomatique de certains (dys)fonctionnements du dictionnaire dit "collaboratif": l'absence de révision systématique des entrées, les décisions prises sans concertation et la collaboration toute relative dans le processus de rédaction des articles contredisent certains poncifs célébrant la "sagesse des foules", l'"intelligence collective" et une certaine "démocratie de la connaissance". Dans le cas qui nous occupe, il n'y a eu, ni pour le domaine du vin, ni pour celui de l'occultisme, aucune discussion collective préalable à une décision qui relève de la ligne éditoriale du dictionnaire. Par ailleurs, aucune initiative de révision systématique des entrées de ces domaines dont le traitement lexicographique est visiblement délicat² n'a été prise alors que "les foules" auraient pu être mobilisées pour, collaborativement, passer en revue les entrées correspondantes et réviser les définitions qui le méritaient, selon des critères discutés conjointement. Rien de tel n'a cependant été entrepris.

Dans ce contexte, la présente étude propose d'interroger le sens de l'adjectif "collaboratif"³ lorsqu'il qualifie un dictionnaire comme le *Wiktionnaire* et de questionner le lien entre (manque de) révisions et (manque de) qualité. Nous nous demanderons tout d'abord si la qualité insatisfaisante des articles est imputable au manque de révisions de ceux-ci. Une questionnement corollaire sera ensuite d'établir si le taux de révision des articles est un prédicteur de la qualité de ceux-ci. Nous examinerons le taux de révision des articles, puis, plus spécifiquement, celui des définitions: les articles (et les définitions) sont-ils souvent révisés et, le cas échéant, le sont-ils par une grande diversité de contributeurs? Nous répondrons à ces questions par une analyse quantitative, à la suite de laquelle nous illustrerons qualitativement les conséquences de révisions trop peu nombreuses, ou inadéquates. Chemin faisant, nous établirons que le lien entre (manque de) révisions et (manque de) qualité passe –notamment– par un déficit de collaboration. À la lueur de ces analyses, nous discuterons, dans un cadre plus large, la nature collaborative du processus rédactionnel et l'opportunité d'évoquer "la sagesse des foules" ou "l'intelligence collective" à propos du dictionnaire contributif.

2. On ne découvre pas ici la complexité du traitement lexicographique des entrées de certains domaines comme celui de l'occultisme ou, plus généralement, qui ont trait à l'existence supposée, revendiquée ou réelle. Dans une section consacrée au « mythe de la neutralité », Atkins & Rundell (2008:427–430) qualifient de « zone grise » ces domaines pour lesquels le choix qui se pose au rédacteur de définition n'est pas entre subjectivité et impartialité, mais plutôt entre plusieurs formes de subjectivité.

3. Nous expliquons en section 7.1 pourquoi nous lui préférons le qualificatif *contributif*, que nous proposons d'utiliser désormais.

2 Méthodes d'évaluation : travaux antérieurs

Dans cette section, nous nous intéressons aux travaux qui se sont attachés à étudier la qualité des projets contributifs (dictionnaire et encyclopédie). Certains de ces travaux consistent en un examen manuel des articles, d'autres optent pour une approche computationnelle. Précisons tout d'abord qu'un *distinguo* s'impose entre les partitions *quantitatif vs qualitatif* et *manuel vs automatique* (ou *computationnel*). On peut mener une étude quantitative manuellement ou de manière computationnelle. En revanche, les études qualitatives comme celles que nous évoquons en section 2.1 sont nécessairement manuelles. On peut produire automatiquement des estimations qualitatives, fondées sur des critères quantitatifs définis *a priori*, pour tenter d'approximer un jugement qualitatif (c'est le cas des mesures de *rigueur* et de *diversité* que nous présentons en section 3.1). On peut également tenter d'établir des liens entre observables quantifiables automatiquement et qualité évaluée manuellement. Ce type d'études quantitatives de la qualité, dont la mise en œuvre est computationnelle, est abordé en section 2.2.⁴

2.1 Études qualitatives manuelles

L'immense majorité des travaux visant l'évaluation des ressources contributives ont porté sur *Wikipédia*. Anthony et al. (2005) notent que l'encyclopédie, du fait de la nature de son processus d'édition et – surtout – de celle de ses contributeurs (qualifiés d'*amateurs*), fait tout à la fois l'objet de louanges et de critiques véhémentes, sans qu'aucune analyse de contenu systématique n'ait été réalisée. Giles (2005) publie la même année dans *Nature* une analyse qualitative qui compare l'encyclopédie contributive à l'encyclopédie *Britannica*. L'étude conclut que la qualité de *Wikipédia* est plus ou moins identique à celle de l'encyclopédie choisie comme référence. C'est le résultat que retiennent et véhiculent encore aujourd'hui les nombreuses citations de l'article, omettant systématiquement de mentionner la réponse de l'éditeur de *Britannica*, qui a relevé ce qu'il considère être de nombreux biais dans la méthodologie adoptée par Giles (Encyclopædia Britannica, 2006). Pour imparfaite qu'elle puisse être, l'étude de Giles (2005), qui visait l'évaluation qualitative⁵ d'un nombre relativement conséquent d'articles pour ce qu'ils sont censés être, *i.e.* un tout cohérent, destiné à la lecture par des utilisateurs humains, n'a pas, à notre connaissance, été reproduite à large échelle depuis.

Concernant le dictionnaire contributif, des lexicographes tels que Hanks (2012), Lew (2014) et Rundell (2017) ont produit des descriptions qualitatives portant sur un nombre très restreint d'articles. Lew, par exemple, fonde sa comparaison entre l'édition anglaise de *Wiktionary* et le *Longman Dictionary of Contemporary English* sur l'étude d'une seule entrée prise au hasard, le verbe *handle*. La taille modeste de leurs observables n'empêche pas ces auteurs de produire des descriptions et des analyses tout à fait intéressantes, mais proscrit toute possibilité de généralisation de leurs observations.

4. Une étude plus approfondie sur les méthodes d'analyse des dictionnaires électroniques est en cours. Elle fera l'objet d'une communication lors de la 4^e édition du colloque *Dictionnaires et culture numérique dans l'espace francophone* en juin 2023.

5. L'étude de Giles est *manuelle* (menée par des experts humains), *qualitative* (les experts identifient pour chaque article ce qu'ils considèrent être des erreurs) et *quantitative* (le nombre d'erreurs par article est comptabilisé sur l'ensemble du corpus pour les deux encyclopédies).

2.2 Études quantitatives computationnelles (de la qualité)

La possibilité d'accéder librement aux ressources contributives a attiré nombre de chercheurs qui ont mené des études quantitatives de l'encyclopédie et, dans une moindre mesure, du dictionnaire. Le problème de ces études quantitatives est qu'elles ne mesurent que ce qui est calculable automatiquement. Or, ce qui est calculable automatiquement n'est pas toujours intéressant, ou n'est pas toujours pertinent pour la description du phénomène qu'il s'agit d'étudier (ici, la qualité des articles). Par exemple, Stvilia et al. (2005) s'intéressent à la « qualité de l'information »⁶ des articles de *Wikipédia*. Ces spécialistes en bibliothéconomie et en sciences de l'information mesurent un certain nombre d'indices qui semblent diversement pertinents, parmi lesquels la longueur des articles (y aurait-il une "bonne" longueur commune à tous les sujets traités ?) ou leur fraîcheur (*currency*), *i.e.* le délai écoulé entre leur dernière mise à jour et la date courante. Mais, dans le cas d'une page qui représente la classification périodique de Mendeleïev ou qui donne les résultats du championnat du monde de football en 1986, l'absence d'une mise à jour récente doit-elle être considérée comme un indicateur de mauvaise qualité ? Un autre critère mis en avant est le degré de formalité de la langue, calculé en comparant les fréquences des parties du discours des articles de *Wikipédia* à celles de la *Columbia Encyclopedia*. À travers cette comparaison, Stvilia et al. concluent à un niveau de langue aussi élevé dans *Wikipédia* que dans la *Columbia Encyclopedia*. Du point de vue opérationnel, il s'agit d'une manière très discutable d'estimer le « degré de formalité » de la langue. Du point de vue théorique, considérer le niveau de formalité de la langue d'une encyclopédie comme un critère de qualité est tout aussi discutable. Plusieurs études, dont celle de Stvilia et al. (2005), mais également celles de Wilkinson & Huberman (2007), Kittur & Kraut (2008) et Daxenberger & Gurevych (2012), tentent d'établir un lien entre diverses mesures, dont la *rigueur* et la *diversité* de Lih (2004), présentées en section 3.1, et la qualité des articles. Ces études n'incluent pas l'évaluation d'un échantillon d'articles par les chercheurs, ou par des experts désignés, comme l'a fait Giles (2005), mais se fondent sur les labels de qualité attribués à certains articles par les contributeurs de *Wikipédia* à travers un processus interne de relecture et de vote.

Concernant le *Wiktionnaire*, ces labels de qualité sont inexploitable. D'une part, on dénombre (au 1/09/2022) seulement 77 "bonnes entrées"⁷ et 4 "très bonnes entrées"⁸. La taille très réduite⁹ de ce jeu de données rendrait statistiquement non significative toute corrélation éventuellement établie entre la qualité des articles et une quelconque mesure. D'autre part, à défaut d'évaluation externe, il conviendrait en premier lieu d'estimer la pertinence de l'attribution par les contributeurs des labels de qualité, ce qui reviendrait à procéder à une étude qualitative manuelle telle que décrite en section 2.1. Enfin, les cri-

6. *Information Quality* est souvent traduit à tort en français par *qualité des données*. Les termes *qualité des données* et *qualité de l'information* existent tous les deux et désignent deux concepts différents. La qualité de l'information est une mesure de la valeur que l'information apporte à l'utilisateur de cette information, ce qui, dans le cas de *Wikipédia*, paraît difficilement quantifiable automatiquement, surtout sans information sur l'utilisateur ni sur le but qu'il poursuit en consultant tel ou tel article.

7. https://fr.wiktionary.org/wiki/Catégorie:Bonnes_entrées_en_français. Le nombre d'entrées portant ce label atteint 86 au printemps 2022.

8. https://fr.wiktionary.org/wiki/Catégorie:Très_bonnes_entrées_en_français

9. Taille très réduite dans l'absolu, mais surtout relativement à celle de la nomenclature du *Wiktionnaire*.

tères de qualité du dictionnaire contributif¹⁰ (qui servent à attribuer les labels *bonne entrée* et *très bonne entrée*) sont très discutables, pour peu qu'ils soient compréhensibles. On lit sous la section « Synonymes » : « Tous les synonymes connus », mais aussi : « quelques-uns suffisent ». Une exhaustivité partielle, en quelque sorte. On lit également sous « Mots apparentés » : « Tous les mots apparentés ». Nous voilà éclairés sur la nature de cette notion. Le critère primordial semble être que toutes les rubriques soient renseignées et, si possible, de manière abondante. Mais l'importance relative des différents éléments de la microstructure n'est pas hiérarchisée. La présence d'une image, par exemple, est recommandée. Elle semble en réalité nécessaire et même plus importante qu'une bonne formulation des définitions¹¹. Concernant ces dernières, on lit dans les critères de qualité qu'elles doivent couvrir tous les domaines (et non tous les sens) et qu'on ne doit pas définir un mot « avec un autre mot de même racine ». Ce qui exclut certains patrons de définitions utilisés couramment pour les adjectifs de couleurs (« de couleur X »)¹², les adjectifs relationnels (« relatif à X », « qui a trait à X », etc.)¹³ et plus généralement toute définition morphologique. Les définitions morphologiques sont pourtant nombreuses dans le *Wiktionnaire*. D'après le dictionnaire contributif, un *islamo-gauchiste* est un « [p]artisan de l'*islamo-gauchisme* » (cf. section 6.1). L'article n'est pas labellisé "bonne entrée". En revanche, l'article *wagonnet*, labellisé, définit le premier sens de la vedette par « Petit wagon [...] » (cf. section 6.2)¹⁴. De même, les sens 1 et 9 de la "bonne entrée" *accommodation* sont définis par « Action d'accommoder [...] » et « Fait d'accommoder [...] ». On lit dans les « [c]onventions pour les définitions¹⁵ » que « [l]es définitions ne débutent pas par la formulation *Variante de...*, les variantes sont indiquées dans une sous-section dédiée et les définitions sont répétées sur chaque page ». Pourtant, le premier sens du deuxième homographe de la "bonne entrée" *astéroïde* est défini par « (*Zoologie*) *Variante de* astéride ». Pour illustrer les définitions, il est conseillé, dans les critères de qualité mentionnés ci-dessus, de « [m]ettre un exemple (au moins) pour chaque sens ». *Couteau*, avec ses 20 exemples illustrant le premier sens, se qualifie haut la main. Mais *corbeau*, malgré l'absence d'exemple pour les 4^e et 5^e sens du nom commun, se qualifie également, comme d'autres "bonnes entrées" dont plusieurs sens sont dépourvus d'exemples.

10. https://fr.wiktionary.org/wiki/Wiktionnaire:Évaluation/Critères_de_qualité (page consultée le 25/05/2022).

11. Au 1/09/2021, toutes les "très bonnes entrées" et 83% des "bonnes entrées" comportent une ou plusieurs images, contre 6% des articles non labellisés correspondant aux formes canoniques (lemmes) des mots lexicaux et 1% pour l'ensemble des formes. Concernant les définitions, on trouve très peu de recommandations dans les différentes pages d'aide du *Wiktionnaire* et dans celles décrivant les critères de qualité des articles. En parcourant les "bonnes entrées", on n'observe pas que leurs définitions soient particulièrement moins problématiques que celles des articles non labellisés (cf. par exemple, en section 6.2, l'article *wagonnet*).

12. Dans le *Wiktionnaire*, l'adjectif *bleu* n'est effectivement pas défini par « de couleur bleue », mais par « De la couleur du ciel en plein jour quand il est dégagé ». En revanche, le premier sens du nom *bleu* est « La couleur bleue ». On observe un traitement identique de nombreuses couleurs.

13. Et pourtant, par exemple, les 4 sens de l'adjectif *fonctionnel* sont tous décrits par une définition morphologique : « Relatif à la fonction », « Qui est bien adapté à sa fonction », « Qui est en état de marche, qui fonctionne » et « (*Physiologie*) Relatif aux fonctions du corps ».

14. Ce qui, en l'occurrence, ne paraît pas aberrant et constitue plus un argument en faveur de l'ajout de nuance dans les critères de qualité qu'en faveur d'une règle absolue de proscription des définitions morphologiques.

15. <https://fr.wiktionary.org/wiki/Wiktionnaire:Définitions> (page consultée le 25/05/2022).

Malgré l'impossibilité de se fonder sur les labels attribués aux articles du dictionnaire contributif pour en estimer la qualité, certaines études quantitatives de *Wiktionary* produisent des éléments de description plus ou moins qualitatifs. Ces études ne sauraient néanmoins tenir lieu d'évaluation. Par exemple, Meyer (2013) donne, dans sa thèse, une description des éditions anglaise et allemande de *Wiktionary* en adoptant « une perspective à la fois TAL et métalexicographique », mais son approche est entièrement quantitative, comme le relèvent Sajous et al. (2020). Meyer compare par exemple les valeurs d'observables quantifiables tels que la couverture lexicale, le degré de polysémie des entrées ou le nombre de marques lexicographiques à celles d'autres ressources¹⁶. Il en ressort quelques tendances très générales : le degré de polysémie des entrées des ressources comparées est plus ou moins corrélé ; *WordNet* affiche une plus grande proportion d'entrées marquées comme relevant des sciences humaines et sociales ; etc. Rien n'est dit, en revanche, sur la pertinence réelle de l'inventaire des sens des unités lexicales ou de la cohérence de l'attribution des marques lexicographiques. Rien n'est dit non plus sur les définitions, qui, pourtant, sont pour Atkins & Rundell (2008:209) au cœur de l'entrée monolingue et en constituent l'élément le plus important. Pour cette raison, et pour celles évoquées en section 3.2, la présente étude porte sur les articles du *Wiktionnaire* dans leur globalité, et plus spécifiquement sur leurs définitions.

3 Rigueur des articles et diversité des contributeurs

Parmi les indicateurs utilisés pour décrire les articles de *Wikipédia*, et par la suite de *Wiktionary*, ou en estimer la qualité, deux mesures – la *rigueur* et la *diversité* – sont couramment utilisées. Nous les décrivons ci-dessous et montrons en quoi elles sont utiles pour la description du dictionnaire contributif. Nous montrerons plus tard pourquoi la première mesure, tout en constituant un élément pertinent de description, nous semble mal nommée.

3.1 Définition de *rigueur* et *diversité*

Dans une étude portant sur ce qu'il appelle le « journalisme participatif », Lih (2004) tente d'évaluer l'impact de la citation des articles de *Wikipédia* dans la presse sur la qualité de ces articles. Si la méthodologie qu'il met en œuvre semble perfectible, Lih est un des premiers à tenter d'estimer quantitativement (et automatiquement) la qualité des articles en introduisant deux mesures, qu'il nomme *rigueur* et *diversité*. La première comptabilise le nombre de révisions d'une page et la seconde le nombre de contributeurs distincts à qui sont dues ces révisions (un même contributeur pouvant réviser plusieurs fois un même article). La mesure intitulée *rigueur* est fondée sur l'hypothèse que « plus de cycles d'édition d'un article permettent un traitement plus approfondi du sujet ou un examen plus minu-

16. Pour mener ses analyses automatiques, Meyer choisit pour comparaison des ressources disponibles au format électronique telles que *WordNet*, mais pas de dictionnaires généraux, ni au format papier, ni au format électronique.

tieux du contenu ». La *diversité* correspond à l'idée qu'« avec plus de rédacteurs, plus de voix s'expriment et apportent autant de points de vue sur un sujet ». Ces deux mesures sont intéressantes en tant qu'éléments descripteurs des articles, ou comme prédicteurs de certains phénomènes. Par exemple, Wolfer & Müller-Spitzer (2016) utilisent la rigueur et la diversité pour décrire la dynamique des éditions anglaise et allemande de *Wiktionary*, tandis que Sajous (2022) les utilise pour détecter les innovations lexicales liées aux phénomènes d'actualité. En revanche, considérer *a priori* que rigueur et diversité sont garantes de la qualité des articles, comme le fait Lih, nous semble constituer un présupposé fort, d'autant que la notion de qualité n'est pas définie formellement. Plusieurs études, comme celle de Kittur & Kraut (2008), proposent de vérifier le lien entre la qualité des articles de *Wikipédia* et un certain nombre de métriques¹⁷ fondées notamment sur les mesures de Lih. La plupart se fondent sur les différents labels de qualité attribués par les utilisateurs de *Wikipédia* à certains articles à travers un processus interne de relecture et de vote. Les métriques sont calculées pour les articles labellisés et pour un échantillon d'articles qui ne le sont pas, afin de vérifier si l'on observe des différences notables de qualité entre les groupes d'articles. Kittur et Kraut notent que le lien observé entre qualité et nombre de contributeurs distincts (*i.e.* la *diversité* de Lih) pourrait être attribué à une causalité inverse. En effet, la qualité des articles ne serait pas nécessairement liée à un nombre important de rédacteurs distincts ; ce serait plutôt la qualité des articles qui aurait tendance à attirer des contributeurs supplémentaires. Par ailleurs, toutes les contributions ne se valent pas et ne devraient pas être prises en compte de la même manière. Daxenberger & Gurevych (2012) établissent une classification des types de contributions dans *Wikipédia* en distinguant les éditions qui visent à faire respecter la politique éditoriale (*e.g.* revenir à une version antérieure après un acte de vandalisme), celles qui opèrent des modifications de surface (*e.g.* formatage, correction orthographique et grammaticale, déplacement de segments textuels) et celles qui ajoutent, suppriment ou modifient le texte de l'article. La quantité d'information (mesurée en nombre de caractères ajoutés, supprimés ou modifiés) est également prise en compte. Cette étude, qui se fonde sur les labels de qualité résultant du processus d'évaluation interne à la *Wikipédia*, ne parvient pas à établir de lien entre qualité et patrons de contribution.

Conçues pour décrire *Wikipédia*, les mesures de Lih sont également applicables au *Wiktionnaire*. Elles ne devraient toutefois pas être surinterprétées, une première précaution à prendre étant de les renommer : leur appellation initiale – surtout pour ce qui concerne la «rigueur¹⁸» – peut en effet orienter les interprétations indûment. Nous montrerons que l'idée que le nombre de révisions d'une page soit nécessairement un indicateur de rigueur est, dans le dictionnaire contributif, un présupposé que les analyses qualitatives démentent. Qualifier la nature des révisions, comme le proposent Daxenberger & Gurevych (2012), est également nécessaire à l'étude du processus d'édition et à l'estimation de la qualité des articles. Il existe de ce point de vue une différence notable entre *Wiktionary* et *Wikipédia* : dans le dictionnaire, contrairement à l'encyclopédie, il est possible de déterminer quels sont les éléments microstructurels édités et d'en tenir compte pour estimer l'influence relative de la révision de ces éléments sur la qualité de l'article. Compléter ou reformuler

17. Une métrique est une combinaison linéaire de plusieurs mesures.

18. Pour notre part, nous parlerons de « nombre (ou taux) de révisions », sauf lorsque nous mentionnerons les travaux de Lih ou ceux qui lui font suite (dans ce cas, nous reprendrons le terme *rigueur*).

une définition contribuera en effet certainement plus à la qualité de l'article correspondant (qu'il s'agisse d'amélioration ou de dégradation) que d'ajouter un 20^e exemple illustrant un sens donné¹⁹ ou un 41^e enregistrement audio illustrant la prononciation d'une entrée²⁰.

3.2 Rigueur et diversité des définitions

Landau (2001:396) écrit, à propos de la première édition d'un dictionnaire, que malgré tout le soin apporté à son élaboration, elle contient toujours de nombreuses erreurs. Concernant plus spécifiquement les définitions des dictionnaires monolingues généraux, Landau écrit qu'au moins deux étapes sont nécessaires à leur rédaction : l'écriture d'un premier jet, puis une révision de la formulation initiale.

In a general monolingual dictionary, defining always requires at least two stages: preliminary or *first-run* defining, and review or final editing [...] No matter who did the first edition, all first-run definitions must be carefully reviewed by senior definers, in no case by the same editor who wrote the original definitions. The review, moreover is not simply critical; the second definer actually redefines the word [...] Only rarely does the preliminary text pass entirely unchanged, more often it is considerably altered, not infrequently entirely rewritten. A definition can be wrong in so many ways and right in so few that it cannot be left to a single writer. (2001:382)

Landau justifie la nécessité des deux passes en expliquant qu'une définition peut être mauvaise de tellement de manières, et correcte de tellement peu, que sa rédaction ne peut être confiée à un rédacteur unique. Ainsi, quel que soit le rédacteur initial, le premier jet doit être révisé par un rédacteur expérimenté et en aucun cas par celui qui a écrit la pre-

19. Voir, par exemple, les 20 exemples du premier sens de l'article *couteau* : <https://fr.wiktionary.org/wiki/couteau> (page consultée le 25/05/2022). Certains articles affichent, pour un sens donné, encore plus de segments textuels présentés comme des exemples (puce et italique suivant une définition), mais qui n'en sont pas (une bonne partie correspondent à des collocations, mais on trouve également des notes d'usage et expressions assorties de marques lexicographiques). Ces articles, comme *avoir* (70 "exemples") et *heureux* (28 "exemples"), sont le résultat d'une mauvaise segmentation d'articles importés automatiquement du DAF8, non révisés bien que visiblement problématiques. Les 20 illustrations du premier sens de *couteau*, en revanche, correspondent bien à des exemples citationnels ajoutés par les contributeurs. On trouve, parmi ces exemples, l'énigmatique « *Tombe 4, couteau-rasoir-précelles en fer* », ajouté sans plus de contexte le 20/12/2016. La référence à une revue d'archéologie est mentionnée, mais il faut toutefois consulter celle-ci, sur le portail *Persée*, pour apprendre que le segment cité dans le dictionnaire correspond à la légende d'une figure représentant un type d'outil trouvé dans la quatrième tombe inspectée lors de fouilles. Notons que l'article *couteau* est labellisé "bonne entrée" le 8/01/2017 (*i.e.* postérieurement à l'ajout du « *couteau-rasoir-précelles* »), ce qui indique soit que l'article n'a pas été relu lors de l'attribution du label, soit que la quantité d'exemples prime sur leur pertinence.

20. D'après le *Wiktionnaire*, *blanc* se prononce blâ au Canada, blā en France (mais blaŋ en français méridional). L'article ne propose pas moins de 41 enregistrements audio dont, notamment, 7 libellés « Paris », 7 autres rattachés à des lieux appartenant à la France non méridionale, 8 de provenances indéterminées et un attribué au « royaume de France ». Même en cas de prononciation uniforme, on peut trouver un nombre déraisonnable d'enregistrements : *ped* se prononce partout pje selon le *Wiktionnaire*, mais l'article propose 30 enregistrements (dont 7 libellés « Paris », 3 étiquetés « France » et 5 « région à préciser »).

mière version de la définition.²¹ Dans le cas du *Wiktionnaire*, on ne peut parler de première (pas plus que de énième) édition du dictionnaire, pris dans sa globalité. Mais l'infrastructure sous-jacente au dictionnaire contributif rend possible l'observation des différentes versions de chaque article individuellement, après chaque modification. Il est également possible – à condition de disposer des moyens techniques et du savoir-faire nécessaire, que nous décrivons en section 4 – d'observer et de quantifier les différentes versions de chaque définition comme autant d'« éditions locales » de cet élément de la microstructure. C'est ce dispositif que nous exploiterons pour vérifier si les préconisations de Landau ont cours dans le *Wiktionnaire*, à savoir si les définitions sont révisées par au moins un contributeur différent du rédacteur initial.

4 Corpus

Nous décrivons dans cette section la nature des données utilisées ainsi que les traitements opérés sur celles-ci afin de catégoriser et quantifier les différents types de révisions.

4.1 Filtrage des données initiales

Le *dump* historique du *Wiktionnaire* est un fichier téléchargeable²², régulièrement mis à disposition par la *Wikimedia Foundation*, qui contient l'historique des révisions de tous les articles du dictionnaire contributif. Après chaque contribution d'un internaute, l'intégralité du contenu de l'article modifié y est stockée, assortie de la date de modification et de l'identifiant de l'utilisateur lorsque celui-ci est enregistré ou de son adresse IP dans le cas contraire (on parle alors d'« utilisateur anonyme »). Le *dump* utilisé dans la présente étude date du 1/09/2021. Stocké dans un fichier qui pèse 67Go, il contient 29 millions de révisions correspondant à 4,5 millions d'articles. Une phase de prétraitement consiste à en éliminer les données qui ne sont pas pertinentes pour la présente étude : pages de discussion et pages utilisateur, pages et sections décrivant des mots autres que français, pages et sections dédiées aux formes flechées, noms propres et mots grammaticaux. Après filtrage, le corpus résultant contient 6 millions de révisions correspondant à 330 000 articles et 13 millions de versions de 497 000 sens.

4.2 Identification et suivi des sens

Lors d'une révision, chaque sens d'un article peut être maintenu à l'identique, modifié ou supprimé. Chacun des sens peut également être déplacé, ce qui conduit potentiellement à un réordonnement de l'ensemble. Un nouveau sens (ou plusieurs) peut également être ajouté à l'inventaire existant à n'importe quelle position (provoquant là encore un réordonnement). Pour déterminer si un sens de l'article révisé a subi ou non une modification,

21. On retrouve ce recours à une écriture à plusieurs mains en programmation informatique. Les « revues de code » consistent, pour un développeur, à détecter les erreurs et les vulnérabilités, mais aussi le manque de lisibilité, dans le code informatique d'un autre développeur. Ces revues se pratiquent parfois « en temps réel », comme dans la méthode *extreme programming* (XP), qui préconise la programmation en binôme (*pair programming*) : un développeur écrit le code, tandis qu'un autre, installé sur le même poste de travail, vérifie le code produit et en détecte les imperfections.

22. <https://dumps.wikimedia.org/>

on ne peut donc préjuger du sens auquel il correspond dans l'article avant modification. Une étape préalable pour établir la correspondance entre l'ensemble des sens de chaque paire de versions successives de chaque article consiste à calculer, pour chaque sens s_i d'une révision R_t , la distance d'édition²³ de Levenshtein (Levenshtein, 1966) avec chaque sens s_j de la révision précédente R_{t-1} . La figure 1 illustre, pour chacun des deux premiers sens de l'article *anis*, modifié le 11/08/2016, la valeur de la distance de Levenshtein par rapport à l'ensemble des sens de la version avant modification, datant du 13/05/2016 (on procède de même pour les sens 3 et 4).

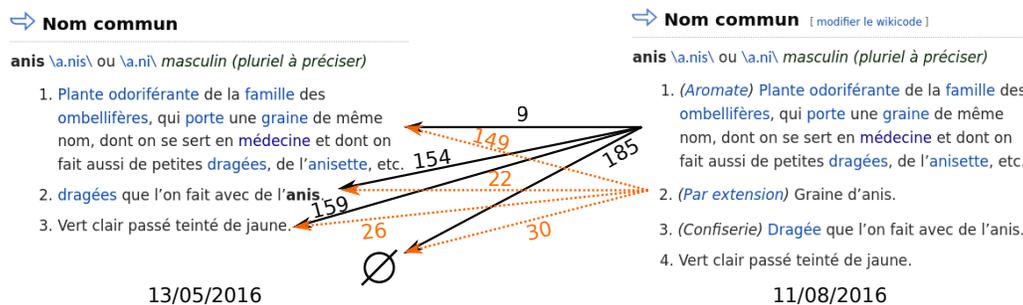


FIGURE 1 : Distances d'édition entre sens avant et après modification de l'article *anis*

Une fois l'ensemble des distances d'édition calculées, l'appariement entre sens des révisions R_t et R_{t-1} est établi en utilisant l'algorithme de Kuhn-Munkres, dit « méthode hongroise » (Kuhn, 1955). Cet appariement, qui établit une correspondance entre chaque sens de la révision courante et un sens donné de la révision précédente ($s_i(t) \rightarrow s_j(t-1)$), minimise la somme des distances d'édition. L'appariement optimal pour la révision de l'article *anis*, représenté en figure 2, correspond au « coût minimal » de $9 + 30 + 14 + 0 = 53$. Il montre que le sens 1 est maintenu à sa place, que le sens 2 est nouveau et que les sens 3 et 4 de la nouvelle version ont été déplacés (ils correspondent respectivement aux sens 2 et 3 de la version précédente). Une fois l'appariement entre sens établi, les sens modifiés sont identifiés par une distance d'édition positive : ici, on voit que les sens 1 et 3 ont été modifiés alors que le sens 4 demeure inchangé.

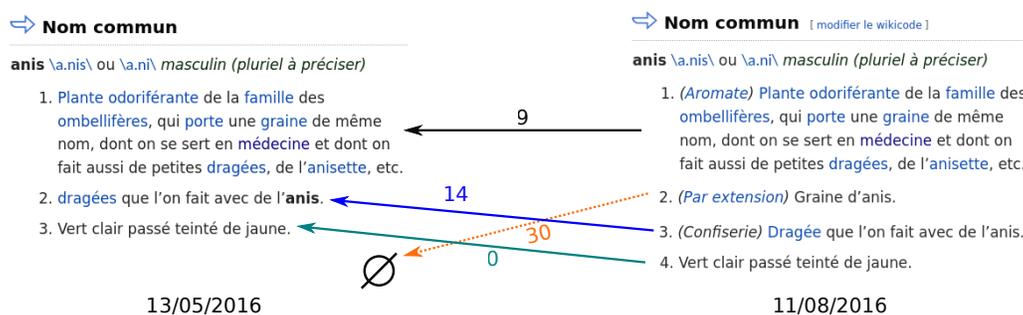


FIGURE 2 : Appariement entre sens de deux révisions successives de l'article *anis*

23. La distance d'édition correspond au nombre minimal d'opérations (insertions, suppressions ou modifications de caractères) nécessaires pour passer d'une chaîne de caractères à une autre.

4.3 Catégorisation des types d'édition

La procédure d'appariement entre sens de deux versions successives d'un article, telle que présentée en section 4.2, permet de distinguer les sens modifiés des sens inchangés selon la valeur de la distance d'édition (positive ou nulle) entre les sens. L'opération suivante consiste à distinguer les modifications surfaciques, *i.e.* celles qui correspondent à l'ajout ou la suppression de tirets, espaces, italique, gras, liens hypertextes ou à un changement de casse (ci-après légendées *déf/surface*) des modifications textuelles (ci-après légendées *déf/texte*), *i.e.* celles qui correspondent à l'ajout, la suppression ou la modification d'au moins un caractère alphanumérique. Nous avons prévu de faire varier le seuil du nombre de caractères modifiés vers des valeurs plus exigeantes afin de distinguer, parmi les modifications textuelles, celles qui sont plus ou moins significatives. Nous verrons en section 5.3 pourquoi cela est inutile.

5 Analyses quantitatives des révisions

Avant d'analyser le taux de révision des pages (section 5.2), puis des sens (section 5.3), nous quantifions en section 5.1 les efforts consentis par "les foules" et étudions leur répartition entre contributeurs.

5.1 Révisions par (types de) contributeurs

Comme pour les autres wikis des projets *Wikimedia*, on distingue dans le *Wiktionnaire* trois types de contributeurs :

- les contributeurs enregistrés, qui ont créé un compte utilisateur et sont identifiés par leur nom d'utilisateur ;
- les contributeurs anonymes, dont seule l'adresse IP est enregistrée lors d'une modification ;
- les robots, *i.e.* des programmes informatiques qui modifient automatiquement les articles. Ces robots effectuent des tâches répétitives portant généralement sur un grand nombre d'articles. Ces tâches peuvent concerner l'import automatique d'articles à partir de sources externes²⁴, des actions de formatage visibles (*e.g.* réorganisation au sein d'une rubrique ou entre rubriques) ou invisibles (*e.g.* modification de l'encodage au format wikicode, à contenu constant) ou encore la génération de formes verbales fléchies²⁵.

La figure 3 représente la proportion de révisions effectuées par types de contributeurs. On observe dans le diagramme en barres empilées de gauche que la majorité (62,64%) des révisions sont effectuées par des robots. Dans la mesure où nous nous intéressons, dans cette étude, à la quantification des définitions révisées par des contributeurs humains, nous considérerons uniquement, sauf mention contraire, les 37,36% des révisions restantes.

24. En 2015, près de 17% des articles faisaient mention d'un import automatique (Sajous & Hathout, 2015).

25. Ce dernier type de modification automatique est absent de notre corpus, qui ne contient que les articles correspondant à des lemmes (*e.g.* nous éliminons de l'article *affluent* la section concernant la flexion verbale et ne conservons que les sections correspondant au nom et à l'adjectif).

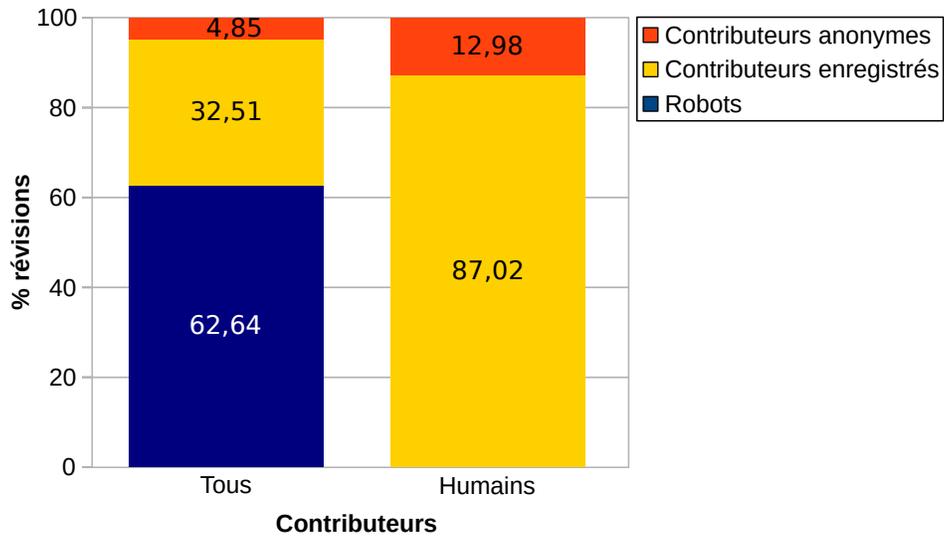
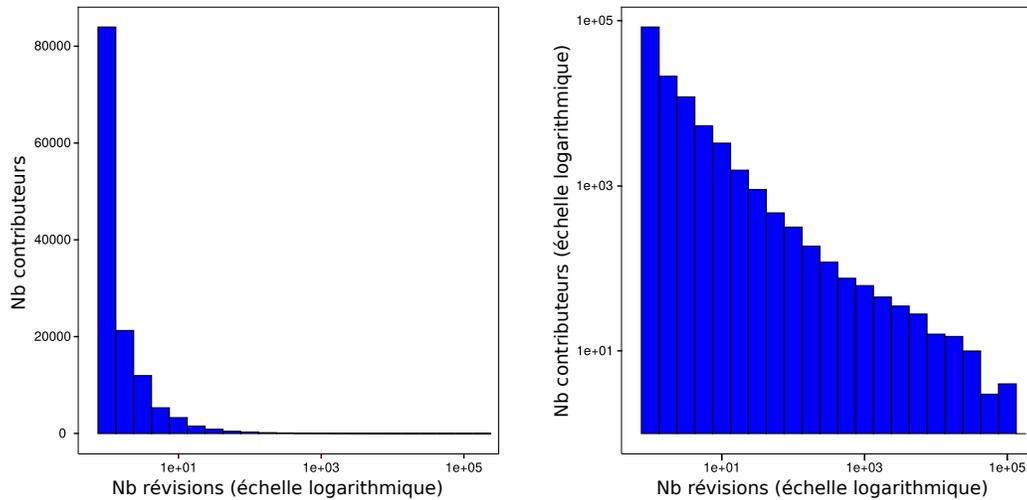


FIGURE 3 : Répartition des révisions par types de contributeurs

Parmi ces dernières (diagramme de droite), près de 13% sont réalisées par des anonymes, le reste étant effectué par des contributeurs enregistrés.

Les histogrammes de la figure 4 représentent tous les deux la distribution du nombre de révisions par contributeur avec un axe des abscisses qui suit une échelle logarithmique. L'axe des ordonnées est représenté selon une échelle décimale dans l'histogramme de gauche, tandis qu'il est représenté selon une échelle logarithmique dans celui de droite. Pour cette distribution et celles dont il est question plus bas, les valeurs minimales, maximales, moyennes et médianes sont données, ainsi que les valeurs des premier et troisième quartiles. Cette distribution, qui suit une loi de puissance inverse, montre que l'immense



Min.	Q1	Médiane	Moyenne	Q3	Max.
1	1	1	17,44	2	140 708

FIGURE 4 : Distribution révisions/contributeurs humains

majorité des contributeurs effectuent peu de révisions (la valeur médiane indique que la moitié des contributeurs n'en ont effectué qu'une seule) tandis qu'une poignée de contributeurs en effectuent un très grand nombre (le maximum étant de plus de 140 000 révisions pour le contributeur le plus actif).

Afin de rendre compte de la proportion de révisions effectuées par ces contributeurs les plus actifs, le tableau 1 indique pour les premiers d'entre eux leur rang, le nombre de révisions effectuées, le pourcentage que ce nombre représente rapporté à l'ensemble de toutes les révisions humaines, ainsi que le pourcentage cumulé des révisions effectuées par les n premiers contributeurs. On voit que les deux contributeurs les plus actifs sont responsables de plus d'une révision sur dix (11,52%), qu'un quart des révisions est imputable à 6 contributeurs, un tiers à 11 contributeurs, la moitié à 27, et les deux tiers à 65 contributeurs.

Contributeurs		Révisions		
Rang	Pseudo	Nombre	% individuels	% cumulés
1	François GOGLINS	140 708	6,22	6,22
2	Diligent	119 754	5,30	11,52
3	Lmaltier	93 283	4,12	15,64
4	Urhixidur	77 563	3,43	19,07
5	JackPotte	77 283	3,42	22,49
6	Jjackoti	53 119	2,35	24,83
...				
11	TAKASUGI Shinji	32 516	1,44	33,67
...				
27	X	17 096	0,76	50,30
...				
65	Coyau	5 324	0,24	66,02
...				

Tableau 1 : Pourcentages individuels et cumulés des révisions effectuées par les contributeurs humains les plus actifs

Le tableau 2 montre que les modifications de définitions se concentrent sur un nombre encore plus restreint de contributeurs : c'est cette fois plus d'une modification sur cinq (21,26%) que l'on doit aux deux premiers contributeurs. Le tiers des modifications de définitions est imputable à 5 contributeurs, la moitié à 15 contributeurs et plus des deux tiers aux 42 premiers contributeurs. Cela confirme les dires d'Anthony et al. (2005) à propos de *Wikipédia* : malgré le "battage" (*hype*) selon lequel une communauté distribuée de participants anonymes peut créer des produits de haute qualité par le biais d'une production *open source*, les études suggèrent que ce type de production est alimentée par un petit nombre d'experts²⁶ qui fournissent la majeure partie du contenu. Le fait qu'une grosse partie de l'activité soit due à un groupe restreint n'est pas surprenant dans ce type de projet contri-

26. Contrairement au cas de projets *open source* comme Linux, où les contributeurs sont réellement des experts en développement informatique, nous transposerons ce qualificatif, dans le cas du *Wiktionnaire*, en contributeur actif.

butif. Cette observation est néanmoins de nature à relativiser la notion de diversité des contributeurs dans le *Wiktionnaire*.

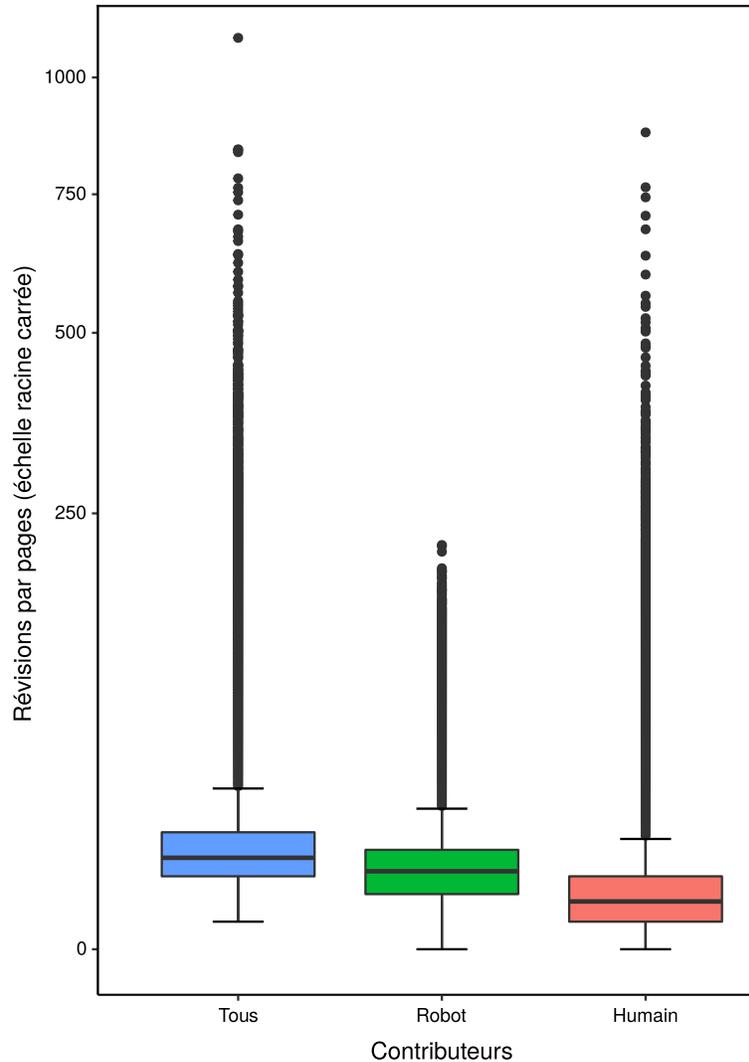
Contributeurs		Révisions de définitions		
Rang	Pseudo	Nombre	% individuels	% cumulés
1	Diligent	23 141	12,96	12,96
2	François GOGLINS	14 828	8,30	21,26
3	Lmaltier	11 097	6,21	27,48
4	Urhixidur	7 222	4,04	31,52
5	TAKASUGI Shinji	4 643	2,60	34,12
...				
15	Béotien lambda	2 053	1,15	50,16
...				
42	X	711	0,40	66,86
...				

Tableau 2 : Pourcentages individuels et cumulés des modifications de définitions effectuées par les contributeurs humains les plus actifs

5.2 Révisions par page

Après avoir étudié dans la section 5.1 la distribution du nombre de révisions par contributeurs, nous proposons un changement de perspective en nous intéressant au nombre de révisions et de contributeurs par page. Les boîtes à moustaches de la figure 5 donnent une représentation des distributions du nombre de révisions par page, réparties selon les types de contributeurs. Le nombre minimal de révisions par page est 1 lorsque l'on prend en compte l'ensemble des contributeurs, car il faut au moins une révision pour créer la page. Pour les distributions ne prenant en compte qu'un type de contributeur, le minimum de 0 correspond, dans la distribution représentant uniquement les contributions humaines, au cas d'une page créée par un robot et jamais révisée, et vice-versa²⁷. Les valeurs extrêmes situées au-dessus des moustaches paraissent nombreuses, mais représentent finalement une poignée de pages (au regard de la taille de la nomenclature) qui attirent un nombre de révisions très supérieur à celui des autres. Les deux articles les plus modifiés sont ceux consacrés à *eau* (878 révisions) et *autrice* (764 révisions). Les deux sont estampillés « très bonne entrée faisant honneur au Wiktionnaire ». Sans commenter la qualité effective de ces pages, notons que si un nombre important de révisions peut contribuer à la bonne qualité d'une page, le fait de proposer une page au vote pour l'obtention d'un label de qualité peut générer un surcroît de contributions jusqu'au vote (en vue de maximiser les chances d'obtention du label). Par ailleurs, une page ainsi labellisée pourrait avoir tendance à attirer

27. L'axe des ordonnées est représenté selon une échelle racine carrée, de manière à ce que les valeurs basses ne soient pas écrasées du fait des nombreuses valeurs extrêmes (représentées par les points au-dessus des boîtes à moustaches). L'habituelle transformation logarithmique est proscrite du fait des valeurs nulles. Omettre ces valeurs nulles ou ajouter 1 à l'ensemble des valeurs, comme il est couramment pratiqué dans ce cas, n'aurait en effet aucun sens ici puisque ces (nombreuses) valeurs nulles (*i.e.* les pages qui n'ont jamais été révisées) constituent le cœur de l'étude.



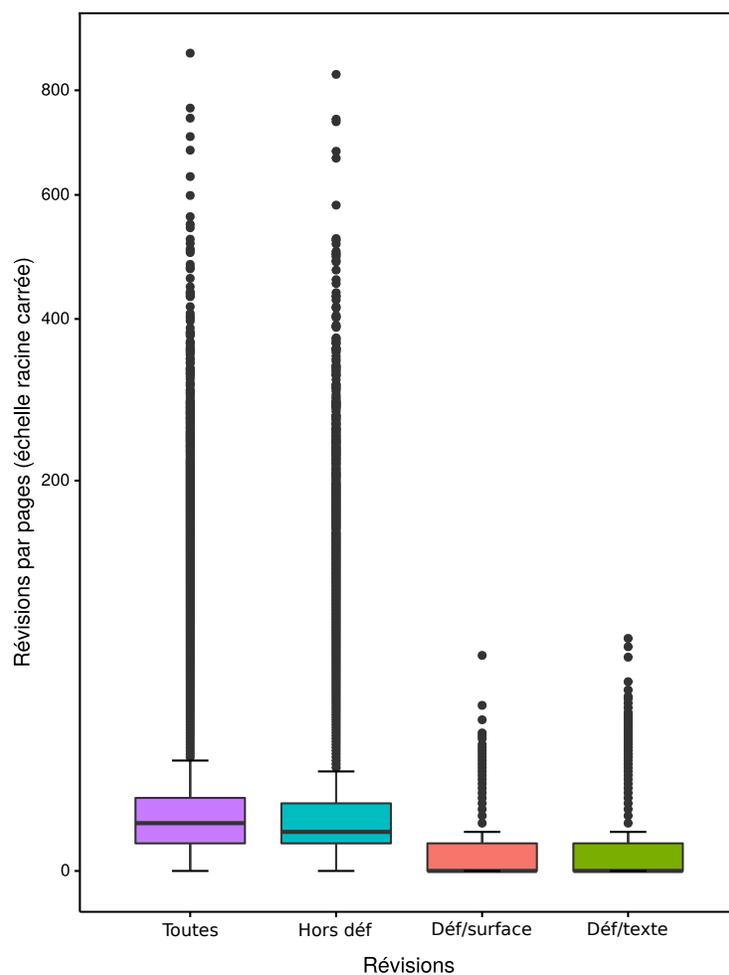
Contributeurs	Min.	Q1	Médiane	Moyenne	Q3	Max.
Tous	1	7	11	18,30	18	1 093
Robots	0	4	8	11,46	13	215
Humains	0	1	3	6,84	7	878

FIGURE 5 : Distributions du nombre de révisions par page, selon le type de contributeur

plus de visiteurs (donc de contributeurs potentiels) comme mentionné en section 3.1. Le nombre moyen de révisions (humaines) par page est de 6,84 mais, étant donné le nombre de valeurs extrêmes, la moyenne donne une vision non prototypique du taux de révision. Avec une valeur de 3 révisions humaines (création comprise), la valeur médiane indique que la moitié des pages ont été modifiées au plus deux fois (après création) par un contributeur.

Les distributions présentées en figure 5 considèrent indistinctement les révisions de tout type, sans prendre en compte la nature des éléments modifiés, au niveau microstructurel. À

l'inverse, la figure 6 donne une représentation des distributions des nombres de révisions (humaines) par article réparties selon qu'elles modifient ou non une (des) définition(s) de l'article et, le cas échéant, selon que les modifications sont superficielles ou textuelles. Les distributions des révisions qui modifient au moins une définition, de manière superficielle ou textuelle, affichent des moyennes proches (autour de 0,5) et une même médiane nulle. L'interprétation est que, pour la moitié des articles du *Wiktionnaire*, le texte des définitions n'a jamais été modifié.

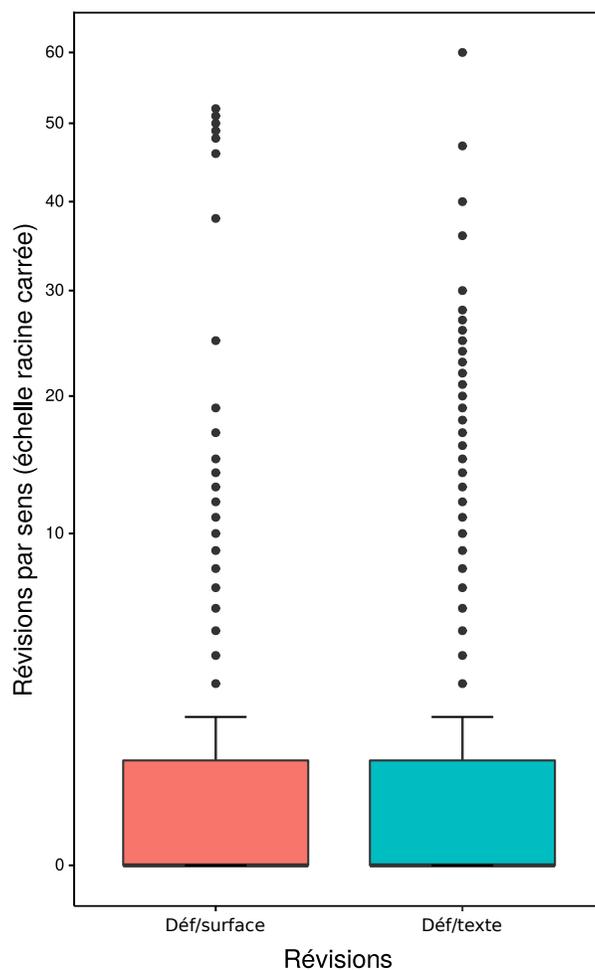


Type d'édition	Min.	Q1	Médiane	Moyenne	Q3	Max.
Toutes	0	1	3	6,84	7	878
Hors déf	0	1	2	5,86	6	833
Déf/surface	0	0	0	0,48	1	61
Déf/texte	0	0	0	0,54	1	71

FIGURE 6 : Distributions des nombres de révisions humaines par page, selon le type d'édition

5.3 Révisions par sens

Nous avons observé en section 5.2 les distributions des nombres de révisions par article, chaque article pouvant contenir plusieurs définitions. Nous menons dans cette section les mêmes observations, mais cette fois sur le taux de révision par sens, pris individuellement. La figure 7 représente les distributions des modifications (humaines) surfaciques et textuelles, par sens, après l'ajout du sens à l'article. Pour les modifications textuelles, la moyenne est de 0,42 et la valeur médiane est nulle. Certains sens étant modifiés plus d'une fois (avec un maximum de 52 pour les modifications surfaciques et 60 pour les modifications textuelles), cela signifie que plus de la moitié des sens ne sont jamais modifiés après leur création par un contributeur humain, quel qu'il soit, *i.e.* rédacteur initial compris.



Type d'édition	Min.	Q1	Médiane	Moyenne	Q3	Max.
Déf/surface	0	0	0	0,37	1	52
Déf/texte	0	0	0	0,42	1	60

FIGURE 7 : Distributions des nombres de révisions humaines par sens, après création, selon le type d'édition

Nous avons vu en section 5.1 que les robots effectuent plus de 62% du nombre total de

révisions. La figure 8, qui représente la répartition du pourcentage de pages et de sens créés par type de contributeur, montre que ces robots créent 37,4% des pages et 30,4% des sens. On pourrait imaginer que les articles et sens importés automatiquement par des robots à partir de sources externes, datant parfois du XIX^e siècle (dans le cas du *Littré*), sont plus révisés (correction d’erreurs dues à l’automatisation, mise au goût du jour des définitions, etc.) que ceux créés par des contributeurs humains, contemporains. La figure 9 montre la répartition des sens révisés après leur création, répartis selon que le “rédacteur” initial, qui crée le sens, est un robot ou un contributeur humain. Cette figure montre que presque 72% des sens importés par les robots (soit 109 550 sens) ne sont jamais révisés.

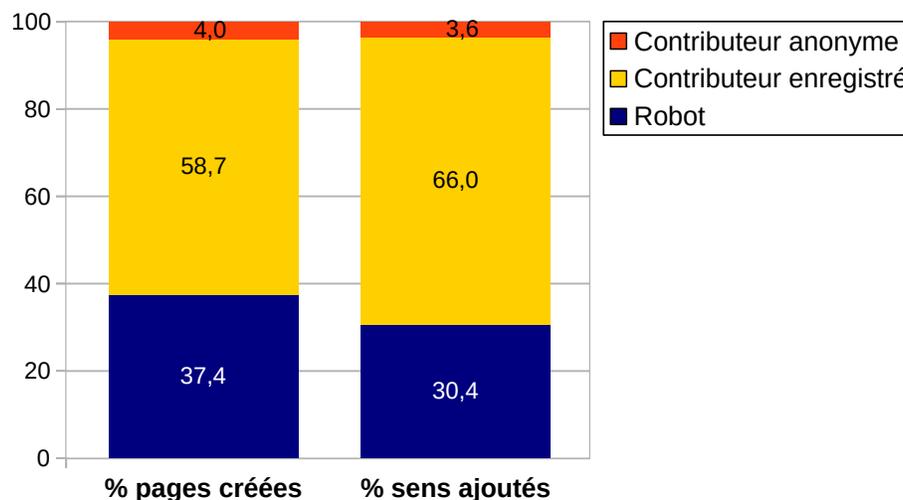


FIGURE 8 : Répartition des créations de pages et de sens par type de contributeur

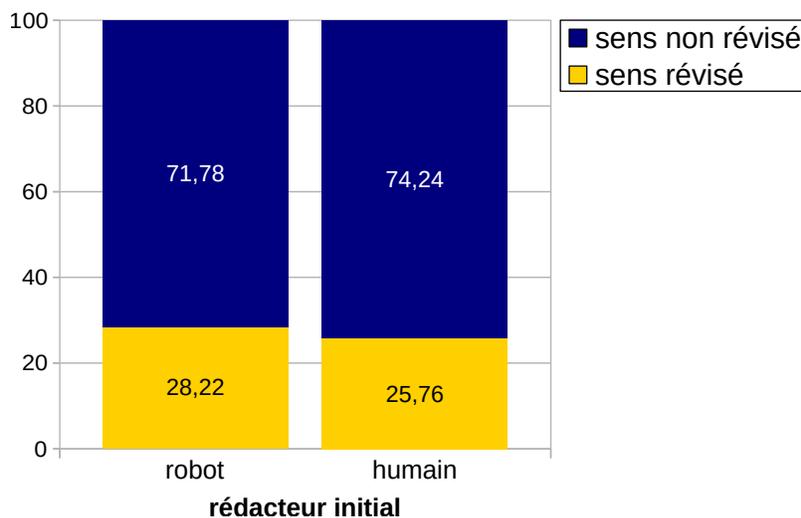


FIGURE 9 : Répartition des sens révisés vs non révisés après leur création, selon le type de “rédacteur” initial

Cette proportion monte à près de trois quarts des sens créés par des contributeurs hu-

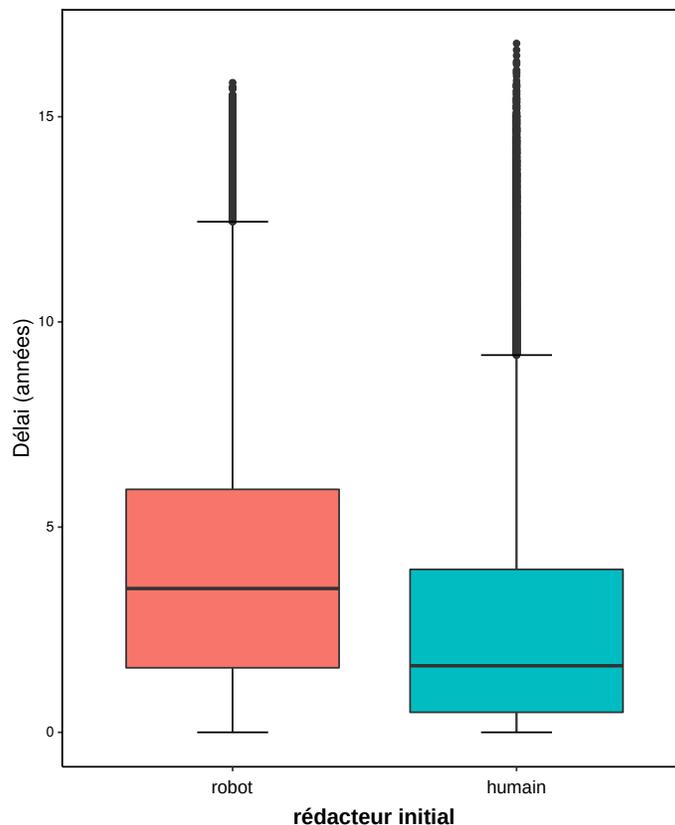
mains (soit 255 318 sens). Les sens importés par des robots n’ont donc que légèrement plus tendance à être révisés que ceux ajoutés par des contributeurs humains. Tous types de “rédacteurs” initiaux confondus, le nombre de sens non révisés après création est de 364 868 sur un total de 496 550, soit 73,5%. Rappelons un élément qui rend ce résultat encore plus frappant : le texte d’un sens est considéré comme ayant été modifié au niveau textuel dès lors qu’un caractère alphanumérique est changé. Si nous augmentions ce seuil peu exigeant – c’est le moins que l’on puisse dire – du nombre de caractères modifiés nécessaire à la comptabilisation d’une modification textuelle, comme nous avons l’intention de le faire pour détecter des modifications plus conséquentes (cf. section 4.3), la proportion de sens non modifiés serait encore plus élevée. Une autre précision qui va dans le même sens : nous n’avons pas cherché à traiter les conflits d’édition et avons comptabilisé, au même titre qu’une contribution “standard”, toute révocation de révision, ainsi que les révocations de révocations, etc. Ne pas comptabiliser ces révocations aurait également augmenté la proportion de sens non révisés.

5.4 Délais de révision

Pour le quart des sens ayant subi une modification, les boîtes à moustaches de la figure 10 représentent le délai écoulé entre l’écriture du premier jet par le rédacteur initial et sa révision par un autre contributeur, selon que le contributeur ayant ajouté le sens est un humain ou un robot.

Un sens rédigé par un contributeur humain sera révisé – s’il l’est – au bout de plus de deux ans et demi en moyenne, avec une valeur médiane d’un an et demi. Si le sens est importé par un robot, et qu’il est révisé, il le sera en moyenne au bout de 4 ans, avec une valeur médiane de 3 ans et demi. Les délais sont longs dans les deux cas, mais on peut s’interroger sur l’urgence visiblement moindre qu’il y a à réviser un sens importé par un robot qu’un sens rédigé par un contributeur humain. Une réponse avancée par le wiktionnariste mentionné dans l’introduction, également présent à Paris lors de la troisième édition du colloque *Dictionnaires et culture numérique dans l’espace francophone*, dans lequel les résultats de la présente étude ont été exposés, est que les définitions de ces sens sont moins relues et révisées parce qu’elles sont importées de sources de confiance et nécessiteraient donc moins de relectures. Rappelons toutefois que les articles de ces “sources de confiance” (DAF8 et *Littré*) datent d’un siècle à un siècle et demi. Pourtant, les définitions ne semblent pas plus devoir être mises en adéquation avec les référents du monde contemporain que le style définitoire ne semble devoir être mis au goût du jour dans le *Wiktionnaire*.²⁸

28. Outre l’obsolescence potentielle du contenu des articles, les imports automatiques posent également des problèmes d’ordre technique. Les programmes utilisés ne traitant pas l’ensemble des formats des articles sources, le contenu de ces derniers est segmenté de manière hasardeuse puis agencé dans l’article cible de manière tout aussi fantaisiste. Nous avons déjà mentionné le cas des “exemples” pour le moins problématiques des articles *avoir* et *heureux* en note de la section 3.1. Les arguments en faveur de la relecture des articles importés ne manquent donc pas.



1 ^{er} rédacteur	Min.	Q1	Médiane	Moyenne	Q3	Max.
Robot	0	1,57	3,50	4,17	5,92	15,83
Humain	0	0,49	1,62	2,70	3,97	16,79

FIGURE 10 : Distribution des délais entre rédaction initiale d'un sens et sa première révision par un autre contributeur

6 Études qualitatives des définitions

Nous avons montré en section 5 qu'à quelques exceptions près, les définitions du *Wiktionnaire* sont très peu révisées. Pour autant, nous n'avons pas encore établi que les définitions peu révisées sont problématiques (ni que d'éventuels problèmes découleraient réciproquement d'un nombre insuffisant de révisions). Or, nul n'est tenu d'accorder une confiance aveugle à Lih, qui postule une corrélation entre nombre de révisions et qualité des articles (cf. section 3.1) ni à Landau, qui affirme la nécessité de réviser systématiquement toute formulation initiale des définitions (cf. section 3.2). On²⁹ pourrait en effet objecter à ces auteurs que, dans le *Wiktionnaire*, les définitions sont correctement formulées dès leur rédaction initiale³⁰. À une telle assertion, nous opposerions ci-dessous quelques contre-

29. Ce *on* pourrait être, par exemple, un wiktionnariste.

30. En citant les préconisations de Landau, nous ne préjugeons d'ailleurs pas de la proportion de projets lexicographiques professionnels qui pratiquent actuellement, ou ont déjà pratiqué, la rédaction des définitions en deux étapes (premier jet, puis révision systématique par un autre rédacteur).

exemples, choisis pour illustrer différentes catégories de problèmes rencontrés dans les définitions du *Wiktionnaire*. La première entrée pêche sur le fond, *i.e.* l'information (en l'occurrence, le manque d'information) contenue dans la définition. La deuxième pose un problème de forme (de formulation). La troisième, enfin, illustre une erreur factuelle indiscutable, décelable par tout utilisateur ou contributeur du dictionnaire contributif, quelle que soit sa compétence lexicographique et linguistique.

6.1 *Islamo-gauchiste* et *islamo-gauchisme* : à entrées polémiques, définitions consensuelles ?

À la suite de Sajous et al. (2019), nous souhaitons examiner des entrées potentiellement polémiques, dont la rédaction des définitions est susceptible de générer des désaccords. Nous souhaitons plus particulièrement examiner des entrées qui font *actuellement* polémique de manière à ce que les articles correspondants ne puissent être importés de sources externes, et observer si “les foules” se mobilisent pour, ensemble, élaborer et mettre au point une description lexicographique qui fasse consensus (la recherche de consensus se traduisant par une succession de révisions).

En France, sous la première mandature d'E. Macron, plusieurs ministres du gouvernement Castex³¹ n'ont eu de cesse – avant de partir en croisade contre le *wokisme* – de fustiger l'*islamo-gauchisme*, phénomène qui « gangrène la société dans son ensemble » auquel « l'université n'est pas imperméable³² » et de vilipender ses responsables : les *islamo-gauchistes*. Dans ce contexte, l'entrée *islamo-gauchiste* semble être un observable remplissant les critères recherchés (un mot polémique au sémantisme fluctuant) et pour lequel on peut s'attendre à observer nombre de révisions. À la lecture de la définition représentée en figure 11, dont on pourrait dire comme Rundell (2017) à propos d'une autre définition de *Wiktionary* « there is so much wrong with this that it is hard to know where to begin », on pourrait penser que l'entrée vient juste d'être créée et qu'elle n'a été encore vérifiée par aucun membre de la communauté des contributeurs (car nul wiktionnariste ne saurait se satisfaire d'une telle incongruité).



FIGURE 11 : *islamo-gauchiste*, 22/04/2021

31. D'abord le ministre de l'Éducation nationale J.-M. Blanquer, à l'automne 2020, puis le ministre de l'Intérieur G. Darmanin et la ministre de l'Enseignement supérieur, de la Recherche et de l'Innovation F. Vidal.

32. F. Vidal, le 14 février 2021 sur la chaîne CNews. L'image de la gangrène (qu'il faut amputer) est régulièrement utilisée dans les discours réactionnaires. Au sujet des « paniques morales » dans lesquelles s'inscrivent – notamment – les discours sur le wokisme et sur l'*islamo-gauchisme*, lire Mahoudeau (2022).

Or, la définition, créée dans le *Wiktionnaire* le 30/12/2014, en même temps que l'article dans lequel elle figure, n'a pas été modifiée jusqu'à la tenue, en octobre 2021, du colloque parisien mentionné en section 1 (la version représentée en figure 11 correspond à la dernière révision avant la tenue du colloque). Pourtant, depuis sa création, l'article a subi 17 révisions humaines réalisées par 10 contributeurs distincts. L'entrée étant monosémique et l'article court, ces contributeurs ne peuvent ignorer la formulation de la définition lorsqu'ils éditent l'article, mais aucun ne semble gêné qu'on puisse être *partisan* d'une religion au lieu d'en être *adepte*, ou qu'on puisse être partisan « de la gauche », au lieu d'être « de gauche », ou adhérent/sympathisant d'un parti qui s'en réclame.

Plus grave, si l'unité lexicale est injurieuse, comme l'indique l'une des deux marques, la définition n'indique pas qui profère cette injure, qui en est la cible, ni en quoi être « partisan à la fois de l'islam et de la gauche », pour peu que cette formule ait un sens, pourrait être infamant³³. Ce type de précision est pourtant nécessaire à la compréhension du sens (dénotation et connotation) de l'entrée par l'utilisateur du dictionnaire³⁴.

Plus grave encore que ces lacunes, notons, en écho aux études de Sajous et al. (2018a, 2019) évoquées dans l'introduction, l'absence de marque de distance dans la définition d'*islamo-gauchiste*, qui vaut prise en charge : sans médiatif dans la définition, le *Wiktionnaire* entérine l'existence – *réelle* et non *supposée* – des *islamo-gauchistes*. La reformulation (du 24/11/2021), après la tenue du colloque organisé en octobre 2021³⁵, de la définition en « Partisan de l'islamo-gauchisme », qui, en renvoyant vers l'article *islamo-gauchisme*, délègue la responsabilité de la description sémantique aux contributeurs de cet article, gomme en partie les problèmes mentionnés plus haut (à supposer que l'*islamo-gauchisme* soit une doctrine ou une mouvance, elle peut bien avoir ses partisans) mais maintient la prise en charge : les *islamo-gauchistes*, comme l'*islamo-gauchisme*, existent bel et bien. Saisir le sens de la nouvelle définition (morphologique³⁶) d'*islamo-gauchiste* impose de comprendre celui d'*islamo-gauchisme*, dont la définition est représentée en figure 12. Or, si le genre prochain choisi pour décrire *islamo-gauchiste* (*partisan*) était mal choisi, que dire de *plateforme politique*, non défini dans le *Wiktionnaire* mais utilisé comme incluant d'*islamo-gauchisme*? On sait seulement que cette « plateforme politique » inclut « certaines idées » (lesquelles?) « progressives » (*progressistes*?) et « certaines idées réactionnaires » (lesquelles?) associées à l'« islam politique », qui, comme

33. Peut-être serait-il plus exact d'écrire que « tout le monde ne voit pas nécessairement » en quoi cela pourrait être infamant.

34. Certains dictionnaires n'hésitent pas à inclure ces renseignements dans leurs définitions. Sajous et al. (2018b) donnent l'exemple du dictionnaire pour apprenants *Macmillan Dictionary* qui précise, dans l'entrée *economic migrant*, qui utilise cette unité lexicale (les gouvernements) et ce qui motive son utilisation (la volonté de distinguer les *migrants* des *réfugiés*). Le rédacteur de la définition trouve ici une juste mesure en fournissant au lecteur certaines explications informatives, sans toutefois commenter les raisons de la volonté des gouvernements d'opérer une distinction entre *migrant* et *réfugié*. Sur ce point, l'absence de commentaire prive le lecteur d'une information utile mais permet la préservation d'une posture relativement neutre, en évitant au dictionnaire d'être accusé de faire un procès d'intention aux gouvernements.

35. Nous avons critiqué cette définition pendant le colloque. Les wiktionnaristes qui ont modifié la définition depuis ne sont pas ceux qui ont assisté à la communication. Mais la communauté des contributeurs actifs étant un petit monde, il n'est pas impossible que des échanges aient eu lieu et que le colloque ait eu une conséquence sur la révision de la définition, comme cela a déjà été le cas avec le colloque de 2018 (cf. section 1).

36. Qui ne respecte donc pas les « critères de qualité » du *Wiktionnaire* (cf. section 2.2).

Étymologie

(2002)^[1] Dérivé de *gauchisme*, avec le préfixe *islamo-*. Terme utilisé pour la première fois par Pierre-André Taguieff dans son essai *La Nouvelle Judéophobie*.

Nom commun

islamo-gauchisme \is.la.mo.ɡɔ.fism\ masculin

Singulier	Pluriel
islamo-gauchisme	islamo-gauchismes
\is.la.mo.ɡɔ.fism\	

1. (Politique) (Injurieux) Plateforme politique incluant certaines idées progressives associées à la gauche et certaines idées réactionnaires associées à l'islam politique.

- Alliance contre-nature d'une partie de la gauche avec un islamisme réactionnaire. — (Sonya Faure, Frantz Durupt, *Islamo-gauchisme, aux origines d'une expression médiatique*, dans *Libération*, 14 avril 2016 [texte intégral])

Note

« Le terme [...] n'a pas d'assise scientifique^[2] ».

Références

- ↑ Shlomo Sand, « Du "judéo-bolchévisme" à "l'islamo-gauchisme" : une même tentative de faire diversion », dans *L'Obs*, 8 juin 2016 [texte intégral]
- ↑ Sonya Faure, Frantz Durupt, « *Islamo-gauchisme, aux origines d'une expression médiatique* », dans *Libération*, 14 avril 2016 [texte intégral]

FIGURE 12 : *islamo-gauchisme*, 11/10/2021 (article tronqué)

plateforme politique, n'est pas défini dans le *Wiktionnaire*³⁷.

On note toutefois deux tentatives visant à apporter de l'information qui pourraient être intéressantes mais peinent à se concrétiser de manière convaincante. La première est la mention, dans l'étymologie, de P.A. Taguieff, à qui le *Wiktionnaire* attribue la première attestation du mot³⁸. Cette précision pourrait permettre de situer le sens actuel de l'unité lexicale par rapport à celui que Taguieff lui attribue lorsqu'il l'utilise en 2002. Un tel développé encyclopédique permettrait également de présenter les critiques que lui opposent ses détracteurs. La seconde précision est la note « le terme [...] n'a pas d'assise scientifique », qui cite un article de *Libération*³⁹ rédigé par deux journalistes du quotidien. Le wiktionnariste responsable de la définition d'*islamo-gauchisme* pourrait trouver dans cet article la matière qui lui permettrait de rédiger une définition pertinente. Un paragraphe qui décrit le sens donné à ce mot en fonction du locuteur qui l'emploie et ce qui motive son utilisation devrait particulièrement retenir l'attention du lexicographe amateur :

37. L'utilisation dans les définitions d'unités polylexicales (*plateforme politique* et *islam politique*) absentes de la nomenclature ne poserait pas problème si le sens de ces unités était compositionnel dans le *Wiktionnaire*, i.e. si le sens d'*islam politique* (resp. *plateforme politique*) pouvait être déduit d'après la définition que donne le dictionnaire contributif du sens d'*islam* (resp. d'un des sens de *plateforme*) et celle d'un des sens de *politique*. Ce qui, à la lecture des définitions, ne s'impose pas de manière patente.

38. On ne doit en réalité à Taguieff que la première attestation écrite du mot.

39. *Islamo-gauchisme, aux origines d'une expression médiatique*, *Libération* (14/04/2016), par Sonya Faure et Frantz Durupt [https://www.liberation.fr/debats/2016/04/14/islamo-gauchisme-aux-origines-d-une-expression-mediatique_1445857/].

Pour ceux qui l'utilisent, l'expression « islamo-gauchisme » est une alerte, un mot « choc » pour décrire l'alliance contre-nature d'une partie de la gauche avec un islamisme réactionnaire. [...] **Pour les autres, ceux qui en font les frais**, elle n'est qu'une arme pour disqualifier une lutte légitime [...] ⁴⁰

On identifie ce paragraphe comme étant la source de la citation représentée en figure 12. Il est déjà saisissant que, pour illustrer une utilisation de l'entrée, le choix de ce qui devrait être un exemple citationnel se porte sur un énoncé (définitoire) qui ne contient pas l'unité lexicale à attester. Par ailleurs, en tronquant la citation (*i.e.* en la privant du médiatif « Pour ceux qui l'utilisent »), le contributeur supprime l'attribution à une source externe et laisse penser que ce sont les journalistes de *Libération* qui définissent (et jugent) l'*islamo-gauchisme* comme une « alliance contre-nature [...] ». Ainsi, tout en passant sous silence, dans la définition, les deux points de vue opposés, l'article du *Wiktionnaire* attribue à tort le premier aux journalistes et prive le lecteur du second, qui lui permettrait pourtant de saisir la variabilité de la connotation en fonction du locuteur.

C'est du même paragraphe que provient la citation (tronquée dans le *Wiktionnaire*) utilisée en note : « Le terme, à vocation médiatique, n'a en tout cas pas d'assise scientifique ⁴¹ ». L'article du quotidien, instructif par ailleurs, pêche ici en énonçant une affirmation gratuite, ou qui apparaît du moins comme telle, faute d'élaborer une argumentation ou d'invoquer une « autorité » qui confirme l'affirmation. Les wiktionnaristes auraient pu, à l'occasion de polémiques qui éclatent à partir de 2020, remplacer la citation de l'article par une citation du même ordre émanant d'un communiqué du CNRS (institution dont on ne doutera pas *a priori* qu'elle puisse faire figure d'« autorité scientifique ») publié en février 2021 qui s'intitule « L'« islamogauchisme » n'est pas une réalité scientifique ⁴² ». Sajous (2022) a montré, en étudiant plus particulièrement le vocabulaire lié à la pandémie de Covid-19, la tendance qu'ont les unités lexicales liées à l'actualité – néologismes ou mots existants – à subir des cycles de révisions relativement importants. Cela n'a pas été le cas ici, malgré le battage médiatique autour de l'expression controversée, que les wiktionnaristes n'ont pu ignorer. La question qui se pose alors est de savoir si aucun d'entre eux n'a eu la curiosité de relire les articles consacrés à *islamo-gauchiste* et *islamo-gauchisme* ou si, le faisant, la qualité de ces derniers leur a paru satisfaisante ⁴³. À ceux-là, nous donnerions

40. C'est nous qui soulignons.

41. Notons que la suppression de « à vocation médiatique » est indiquée par les crochets « [...] » alors que celle de « en tout cas » n'est pas indiquée.

42. <https://www.cnrs.fr/fr/1-islamogauchisme-nest-pas-une-realite-scientifique>
Le communiqué publié par le bureau de presse du CNRS le 17/02/2021 fait suite à la demande de la Ministre de l'enseignement supérieur, de la recherche et de l'innovation de mener une étude visant à apporter un éclairage scientifique sur certains champs de recherche, notamment les études postcoloniales. Dans son communiqué, le CNRS affirme que l'islamogauchisme ne correspond à aucune réalité scientifique, que les contours du terme sont mal définis, et condamne avec fermeté la remise en cause de la liberté académique et en particulier les tentatives de délégitimation de différents champs de la recherche. Un article du quotidien *Le Monde* daté du 29/03/2023 révèle que, contrairement à ses dires, la ministre Frédérique Vidal n'a jamais diligenté l'enquête annoncée en février 2021.

43. Au printemps 2022, la définition d'*islamo-gauchiste* est toujours « Partisan de l'islamo-gauchisme ». Celle d'*islamo-gauchisme*, modifiée le 24/11/2021, devient « (Politique) (Injurieux) Expression politique et polémique cherchant à stigmatiser certaines idées progressives associées à la gauche en les présentant comme étant associées à l'islam politique ». Elle reste très perfectible (cf. notamment l'adjectif *progressive*, l'anthropomorphisme qui prête à « l'expression politique » la volonté de « chercher à stigmatiser » et l'*islam politique*, toujours absent de la nomenclature), mais sa révision a le mérite d'aller dans le bon sens.

deux conseils de lecture. Le premier est celui de Vincent (2022) qui préconise un traitement polyphonique des entrées polémiques et illustre son propos par l'étude du traitement de *woke*⁴⁴ par plusieurs dictionnaires. Le second conseil est la lecture de l'article du dictionnaire *Orthodidacte* dédié à *islamo-gauchisme*⁴⁵, qui, justement, offre une description polyphonique, et, par ailleurs, montre que les définitions aristotéliennes ne sont pas toujours les plus adaptées au traitement de certaines entrées. Une définition phrastique avec un développé encyclopédique ou, selon le type de dictionnaire et le public visé, une notice culturelle, semble être une bonne alternative. Une autre possibilité est de compléter la définition aristotélienne par divers types de remarques à travers lesquelles « le lexicographe s'adresse à l'utilisateur » (Vincent, 2021). Ce que font parfois les contributeurs du *Wiktionnaire* (cf. la note sur l'absence d'assise scientifique du terme *islamo-gauchisme*), avec un succès variable. Reste alors à déterminer, en premier lieu, le genre prochain le plus adéquat. Renoncer à un nom d'agent au profit d'une expression métalinguistique, suivie de la caractérisation des locuteurs qui emploient l'entrée définie, permettrait d'éviter la prise en charge. Pour *islamo-gauchiste*, l'incluant pourrait être *insulte* (au lieu de *partisan*), que complèteraient la mention des utilisateurs de l'unité lexicale, de leur but et de leur cible : « utilisée par... », « à propos de... », « afin de... »⁴⁶.

6.2 Wagonnet : définition vs explication d'un terme spécialisé

Afin d'étudier la définition d'une entrée moins polémique que celle analysée en section 6.1, nous proposons d'examiner celle d'un terme spécialisé. L'article *wagonnet* est créé dans le *Wiktionnaire* le 22/02/2006 à partir d'un import du DAF8, qui fournit la définition du sens général, demeurée inchangée jusqu'à aujourd'hui (printemps 2022) : « Petit wagon fait pour être traîné ou poussé [...] ». Le sens héraldique, représenté en figure 13, est ajouté le 24/04/2019 et demeure lui aussi inchangé dans la dernière révision de l'article (9/08/2021). On doit cette définition au spécialiste de l'héraldique du *Wiktionnaire*, qui a contribué à la plupart, si ce n'est la totalité, des définitions des entrées de ce domaine. On peut supposer que ce contributeur a rédigé la définition lui-même, mais le doute est permis⁴⁷. Le rédacteur de cette définition à visée terminologique, assortie d'un développé encyclopédique, n'a pas fait le travail d'adaptation nécessaire à la compréhension par un lectorat non spécialiste. Vulgariser de manière satisfaisante une définition terminologique,

44. Outre l'illustration convaincante de l'intérêt d'un traitement polyphonique pour *woke*, l'étude de cas de Vincent comporte une analyse du traitement de ce mot opéré par le *Wiktionnaire*, que le lecteur pourra comparer avec les traitements d'*islamo-gauchiste* et d'*islamo-gauchisme* présentés ici. En France, *woke*, *wokisme* et *wokiste* succèdent à *islamo-gauchisme* et *islamo-gauchiste* qui, bien qu'ayant des sens différents, sont utilisés (à tort) de manière interchangeable par les mêmes locuteurs pour dénigrer les mêmes cibles.

45. <https://dictionnaire.orthodidacte.com/article/definition-islamo-gauchisme> (page consultée le 25/05/2022)

46. Cette stratégie est d'ailleurs – rarement – mise en œuvre dans le *Wiktionnaire*. Par exemple, *bloke* est défini comme une « Insulte utilisée par les Canadiens français pour désigner les Anglais, en particulier les Canadiens anglais ».

47. On ne trouve ni référence à une source externe dans l'article, ni trace de cette définition ailleurs sur le web, ni description du sens spécialisé dans le DAF8, le *Littré*, le *Trésor de la langue française*, le *Petit Larousse*, le *Petit Robert* et *Usito*. Mais le contributeur ne précise pas non plus la source de l'exemple citationnel attestant l'usage héraldique de *wagonnet*, alors qu'une recherche sur le web indique qu'il s'agit de la description du blason de la commune de Cosnes-et-Romain, reprise dans une demi-douzaine de pages (sans qu'on puisse pour autant trouver mention de la provenance originelle).

2. (*Héraldique*) **Meuble** représentant l'engin du même nom dans les **armoiries**. Il est généralement représenté sous la forme d'une **berline de mine** contenant ou non une substance éventuellement d'une autre couleur. Il est présenté **de profil** et **posé en fasce** sans qu'il y ait de distinction **dextre/senestre**. À rapprocher de **locomotive** et **TGV**.

- *Écartelé au 1° de sinople à une tierce ondée alaisée d'argent, au 2° de pourpre à l'épi de blé d'or en pal, au 3° d'azur à la hache d'argent en barre fendant une souche de sable, au 3° d'or au **wagonnet** de sable enflammé de gueules. Sur le tout semé de croix recroisetées au pied fiché d'or à deux bars adossés de même, qui est de la commune de Cosnes-et-Romain de Meurthe-et-Moselle → voir illustration « armoiries avec un wagonnet »*

FIGURE 13 : *wagonnet* (sens 2), 9/08/2021

on le sait, peut s'avérer délicat.⁴⁸ Entre autres difficultés, il s'agit souvent de choisir entre définir et expliquer. Une définition suffisamment précise pour satisfaire un spécialiste peut se révéler incompréhensible au plus grand nombre, alors qu'une explication pédagogique risque d'être inexacte, voire fautive. Mais, selon les disciplines et – au sein d'un champ donné – selon les concepts à définir, les degrés de technicité sont très variables. On voit mal comment la concision habituelle d'une description lexicographique permettrait de définir, par exemple, une structure algébrique particulière (*e.g. anneau, corps, groupe, etc.*) à un lecteur ne possédant pas un arrière-plan universitaire en mathématiques.⁴⁹ Pour confidentiel que soit le domaine de l'héraldique, la compréhension de ce qu'est un wagonnet dans les armoiries n'est en revanche pas hors de portée du lecteur profane. Quant à la rédaction de la définition correspondante, si le rédacteur initial, connaisseur du domaine, n'est pas conscient des problèmes que pose sa formulation, il semble possible qu'un autre contributeur du *Wiktionnaire* puisse l'améliorer sans avoir de connaissance du domaine.

On note, dans la formulation de la définition, une lourdeur due à une forte redondance (« Meuble représentant [...] Il est généralement représenté [...] Il est présenté »)⁵⁰, mais également par endroits un manque d'explicite. Par exemple, rien ne dit que *meuble* n'est pas employé ici dans son sens le plus général, d'autant que le lien hypertexte qui porte sur ce terme renvoie à l'ensemble de l'article correspondant et non au sens concernant spécifiquement l'héraldique (figurant en quatrième position). À propos de la substance contenue (ou non) dans la berline de mine, la formulation « d'une autre couleur » est également implicite : on suppose – mais on suppose seulement – que la substance est d'une

48. Pour ne citer que quelques-uns des nombreux travaux sur le sujet, lire par exemple Boulanger & L'Homme (1991), Béjoint (1998) et Becker (2015).

49. Pour ces entrées, le *Petit Robert* fait le choix de définir (de façon discutable), non de vulgariser. Le *Wiktionnaire* opte pour des définitions spécialisées pour *anneau* et *corps*. Pour *groupe*, sans tenter une explication grand public (ce qui paraît sage), le *Wiktionnaire* donne une définition plus générale : « (*Mathématiques*) Structure algébrique abstraite munie de certaines propriétés mathématiques. ». Cette dernière option ne paraît pas la plus mauvaise, à supposer qu'un sens aussi spécialisé et aussi peu répandu dans la langue générale doive être décrit dans un dictionnaire de langue. Mais on peut se demander d'une part s'il existe une chance qu'un non spécialiste ait affaire à une telle acception (et qu'il cherche sa définition dans un dictionnaire) et, d'autre part, pourquoi un spécialiste chercherait la définition d'un terme dans un dictionnaire de langue plutôt que dans un ouvrage spécialisé.

50. Sans parler des reprises par le pronom anaphorique qui peut aussi bien renvoyer à *wagonnet*, *meuble* ou *engin du même nom*.

autre couleur *que la berline de mine*. Les termes spécialisés devraient être remplacés par des termes plus simples lorsque cela est possible, ou du moins paraphrasés. Par exemple, la proximité orthographique de *en fasce* avec *en face* pourrait laisser croire à une altération graphique et une parenté sémantique, mais il faut suivre deux liens pour lire qu'une *fasce* est « [u]ne des pièces honorable de l'écu [...] ». Une alternative à la paraphrase pourrait être la mise en place d'infobulles qui afficheraient la glose des termes spécialisés. D'après les wiktionnaristes présents au colloque parisien (cf. section 1), la communauté des contributeurs aurait voté en défaveur de l'utilisation des infobulles dans le *Wiktionnaire*, au motif qu'elles surchargeraient les pages.⁵¹ La préférence donnée à des liens hypertextes plutôt qu'à une définition *in extenso* est un choix discutable, surtout lorsque ces liens sont utilisés de façon aussi critiquable que pour *berline de mine*. Le lien portant sur le syntagme entier laisse penser qu'un clic permettra à un internaute de consulter la description de l'unité polylexicale spécialisée, mais il renvoie seulement à l'article *berline*, dans lequel on ne trouve mention d'aucune *berline de mine*. Dans le troisième sens de *berline*, un lecteur coopératif trouvera en revanche mention d'un « wagonnet sur rails utilisé pour le transport du minerai du fond de la mine ». Il en déduira finalement la compositionnalité du sens de *berline de mine*. Quant à l'utilisation de *dextre* et *senestre*, utilisés (notamment) en héraldique, la définition perdrait-elle en exactitude, en précision ou en pouvoir explicatif, à utiliser *droite* et *gauche*⁵² ?

Depuis l'ajout du sens héraldique, l'article a été révisé par trois autres contributeurs, qui ont ajouté une prononciation et deux traductions. Rien ne permet de savoir s'ils ont relu la définition du sens spécialisé. Personne, en tout cas, n'a jugé bon d'en modifier la formulation, à laquelle s'applique pourtant une critique de Hanks (2012) : « as in many traditional dictionaries, the definitions may succeed in defining, but they do not do a very good job of explaining ». Dans le cas présent, la deuxième assertion est indiscutable. La première, en revanche, n'est pas certaine. Pourtant, l'article consacré à *wagonnet* arbore un label de qualité : un bandeau situé en bas de page nous informe que « l'entrée en français a été sélectionnée comme bonne entrée ». L'attribution de ce label par la communauté des contributeurs interroge : soit la formulation des définitions n'est pas un critère de qualité dans le *Wiktionnaire* (ce qui confirmerait l'impossibilité, mentionnée en section 2.2, d'exploiter ces labels de qualité), soit les contributeurs de ce dictionnaire ont une conception toute particulière de ce que devrait être une bonne définition (les deux propositions n'étant pas exclusives).

51. Cette justification est pour le moins déroutante : une infobulle n'apparaît dans une page web que lorsque l'utilisateur survole une zone à laquelle est associée cette infobulle. Il suffit de laisser le pointeur de la souris hors de la zone de texte pour qu'aucune infobulle ne vienne « surcharger » la page. Notons par ailleurs que les infobulles sont abondamment utilisées dans l'encyclopédie *Wikipédia* pour afficher un bref résumé des termes sans avoir besoin de quitter la page consultée. La communauté des contributeurs de *Wikipédia* (plus nombreuse que celle du *Wiktionnaire*) ne semble pas considérer que cette utilisation systématique des infobulles génère une quelconque surcharge.

52. De la même manière, le premier sens de *berline* est défini par « Sorte de grande voiture hippomobile » : en quoi *hippomobile* est-il un meilleur choix que « tractée par des chevaux » ? Ou, à défaut de remplacer le terme par une paraphrase (si le dictionnaire se donne pour mission d'enrichir le lexique de son utilisateur), pourquoi ne pas faire suivre le terme de cette paraphrase, éventuellement parenthésée (ou opter pour l'ordre inverse : paraphrase suivie du terme entre parenthèses), d'autant que la contrainte de place est inexistante dans le *Wiktionnaire* ?

6.3 Clavardeuse : mélange des genres

Les traitements lexicographiques présentés en sections 6.1 et 6.2 se caractérisent par une incapacité à remplir leur fonction descriptive, une formulation maladroite ou une inadaptation aux utilisateurs potentiels⁵³. Nous présentons ci-dessous un article comportant une erreur factuelle et objective, que tout relecteur natif, même ignorant du référent décrit et dépourvu d’expertise linguistique et lexicographique, peut déceler. Le mot *clavardeuse* qui désigne, au Canada, une femme qui clavarde, *i.e.* impliquée dans un *chat* (messagerie instantanée), fait son entrée dans le *Wiktionnaire* le 26/07/2009, dans un article tel que représenté en figure 14.

	Singulier	Pluriel
Masculin	clavardeur \kla.vaʁ.dœʁ\	clavardeurs \kla.vaʁ.dœʁ\
Féminin	clavardeuse \kla.vaʁ.døz\	clavardeuses \kla.vaʁ.døz\

FIGURE 14 : *clavardeuse*, 26/07/2009

Nous laisserons de côté le fait que le nom d’agent féminin soit défini comme le féminin du nom d’agent masculin équivalent – c’était, à l’époque, le traitement de rigueur dans le *Wiktionnaire*, qui a évolué depuis sur ce point⁵⁴. Plus flagrante est la mention du genre masculin, dans ce qui s’appelle dans *Wiktionnaire* la « ligne de forme⁵⁵ », qui contredit le genre féminin de la définition formulée à la ligne juste en dessous. Déceler cette erreur n’exige d’un locuteur natif ni une formation lexicographique pointue, ni une intuition linguistique affûtée, mais il aura fallu attendre presque sept ans pour que cette erreur soit corrigée, le 17/05/2016 (cf. figure 15), à l’occasion d’un remaniement salutaire⁵⁶ des noms d’agents féminins.

53. À notre connaissance, le *Wiktionnaire* ne fait nulle part mention du public visé. Gageons qu’il s’agit du plus grand nombre.

54. Du moins partiellement, l’application d’un changement de traitement n’étant jamais totalement systématique dans le *Wiktionnaire*. Ainsi, au printemps 2022, le nom *fricasseuse* est toujours le féminin de *fricasseur*, *bourlingueuse* celui de *bourlingueur*, etc.

55. La *ligne de forme* est constituée de la vedette, éventuellement suivie de sa prononciation et d’informations grammaticales : https://fr.wiktionary.org/wiki/Aide:Jargon#ligne_de_forme

56. Salutaire, le remaniement ne l’est cependant pas lorsque les wiktionnaristes féminisent artificiellement certains noms d’agents masculins – dont l’inclusion dans la nomenclature est discutable – pour produire des noms féminins qui ne sont pas attestés. Par exemple, *rampeuse* devient dans le *Wiktionnaire* une « Femme qui rampe », après avoir été le « Féminin singulier de rampeur », l’article actuel précisant « Exemple d’utilisation manquant ». Le nom masculin *rampeur*, « Homme qui rampe », est illustré par un exemple citationnel tiré d’un ouvrage publié en 1952 dans lequel des soldats « arrivent en rampant ». Il n’est pas improbable que l’attestation soit une création d’auteur ponctuelle. Pour ce qui est de la forme au féminin, une recherche sur Internet ne renvoie que 3 occurrences du nom *rampeuse* utilisé pour désigner un agent humain. Dans ce cas,

⇒ Nom commun	
clavardeuse \kla.vaʁ.døz\ féminin (équivalent masculin : <i>clavardeur</i>)	
1. Femme, fille qui <i>clavarde</i> .	
Singulier	Pluriel
clavardeuse	clavardeuses
\kla.vaʁ.døz\	

FIGURE 15 : *clavardeuse*, 17/05/2016

Le cas de *clavardeuse* n'est pas isolé : 176 entrées ont été concernées par ce mélange des genres. En consultant, dans l'historique, les commentaires laissés lors de la création des articles, on constate que 141 de ces 176 erreurs ne sont pas dues à un phénomène de syncrétisme collectif, mais à l'utilisation par plusieurs contributeurs d'un même « gadget » (script) qui « permet d'attirer l'attention sur les flexions manquantes en français, et aide leur création en proposant un wikicode⁵⁷ ». L'erreur récurrente résulte donc soit d'une erreur de conception du script, soit d'une mauvaise utilisation de celui-ci. Au printemps 2022, il reste encore 10 entrées, sur les 176 initialement concernées, dans lesquelles le problème persiste. Par exemple, dans la page *friteuse*, la deuxième entrée est étiquetée *masculin* et définie comme « Féminin singulier de friteur » (un *friteur* étant, d'après le *Wiktionnaire*, un « Cuisinier spécialement chargé des fritures »). Quoi qu'il en soit, si une erreur aussi factuelle peut persister aussi longtemps que pour *clavardeuse*, c'est bien faute de processus de relecture des articles.

7 Synthèse et discussion

Comme nous l'avons mentionné dans l'introduction, la présente étude trouve son origine dans un travail antérieur, mené par Sajous et al. (2019), qui formulaient deux constats. Premièrement, un constat d'incohérence du *Wiktionnaire* en matière de respect de (ou d'entrave à) son principe de neutralité de point de vue, dû pour partie à l'absence de processus systématique de révisions. Deuxièmement, un constat d'absence de discussion dans la communauté des wiktionnaristes sur des décisions concernant la ligne éditoriale du *Wiktionnaire*. Partant du premier constat, nous nous sommes interrogé sur le taux de révision des articles, et sur les conséquences possibles d'un faible taux de révision sur la qualité des définitions. Partant du second, nous avons questionné la nature collaborative du processus de rédaction du dictionnaire.

L'étude quantitative de la section 5 a montré que les contributeurs actifs du *Wiktionnaire*

il désigne toujours un bébé. Pour le reste, on ne trouve qu'une poignée de faux positifs (*rappeuse*, *rampe + use*), d'emplois adjectivaux (*excavatrice*, *porte*, *pelle rampeuse*) et de locutions (*rampeuse de nuit* est le nom d'une araignée). On note que, contrairement à ce qu'énonce la page *Critères d'acceptabilité des entrées* (« il faut vérifier que le mot est bien attesté »), l'absence d'attestation ne semble pas être, dans la pratique, un obstacle à l'inclusion. Ce renoncement à suivre la ligne éditoriale paraît même totalement assumé : dans l'article créé le 13/12/2018, *saucisseuse* est définie comme « Féminin singulier de saucisseur ». Le 5/05/2020, la définition initiale devient « (*Désuet*) Ancien nom des charcutières. », en même temps qu'apparaît un second sens (« Machine servant à fabriquer des saucisses. »). La nouvelle définition du premier sens est suivie d'une note qui stipule qu'« [a]ucune attestation écrite n'a pu être trouvée pour cet usage ».

57. <https://fr.wiktionary.org/wiki/Aide:Gadget-CreerFlexionFr>

sont peu nombreux, que les articles sont peu révisés, et que les définitions le sont très peu : au moment de l'étude, plus de 73% des définitions demeurent à l'état de « premier jet ».

L'étude qualitative menée en section 6 a permis de décrire des problèmes de différentes natures qui apparaissent dans les définitions du dictionnaire contributif. En nous appuyant sur ces analyses, nous apportons ci-dessous les réponses aux questions que nous posions en introduction. Le faible taux de révision nous conduit à relativiser, en section 7.1, la nature collaborative du processus de rédaction. Devant ce niveau de collaboration limité et la diversité restreinte des contributeurs actifs, nous nous interrogeons plus généralement, en section 7.2, sur la notion d'« intelligence collective » ou de « sagesse des foules » dans le dictionnaire contributif. Nous concluons, en section 7.3, que, parmi les liens qui, dans le *Wiktionnaire*, mènent du manque de révisions du contenu existant au manque de qualité, un des maillons faibles est certainement le déficit de collaboration.

Avant de poursuivre, nous rappelons, de manière synthétique, quelques éléments mis en évidence dans cette étude⁵⁸ :

1. Dans le *Wiktionnaire*, la diversité potentielle des contributeurs est très relative :
 - un contributeur sur deux n'a effectué qu'une seule révision d'article ;
 - 11 contributeurs ont réalisé un tiers de l'ensemble des révisions et 27 contributeurs en ont réalisé la moitié ;
 - la diversité est encore moindre concernant la révision des définitions :
 - 5 contributeurs sont responsables du tiers des révisions de définitions ;
 - 15 sont responsables de la moitié de celles-ci.
2. Les articles sont peu révisés, les définitions le sont très peu :
 - la moitié des pages ont été révisées moins de 3 fois par un contributeur humain ;
 - près de trois quarts des définitions n'ont jamais été révisées.
3. Pour le quart des définitions qui sont révisées, le délai entre création et révision est long :
 - plus d'un an et demi pour les définitions créées par des contributeurs humains ;
 - trois ans et demi pour les définitions importées automatiquement.
4. Un nombre élevé de révisions (ou de contributeurs distincts) d'une page n'est pas un prédicteur de qualité.

58. L'ensemble de ces rappels se fonde sur les valeurs exactes ou médianes des phénomènes comptabilisés au 1/09/2021. Rappelons que nos choix nous conduisent à ne pas présenter les résultats les plus spectaculaires : étude des articles concernant uniquement les formes canoniques (les formes fléchies sont encore moins révisées), seuil d'un unique caractère modifié pour comptabiliser une définition révisée, comptabilisation des révocations de révisions.

7.1 Dictionnaire collaboratif vs patchwork lexicographique

Une caractéristique souvent associée aux projets contributifs (lexicographiques et encyclopédiques) est la capacité des contributeurs à réviser (reformuler, corriger, amender) le contenu des articles. Ce que les wiktionnaristes francophones⁵⁹ font très peu, y compris dans le cas d'entrées à caractère polémique, pour lesquelles une réelle collaboration conduirait idéalement à une définition à la fois pertinente et consensuelle. Mais, dans le *Wiktionnaire*, c'est souvent, pendant de longues périodes et pour de nombreuses définitions, le premier qui a parlé qui a raison (du fait qu'il est souvent également le dernier). Une fois une entrée définie, les contributeurs suivants complètent l'article en ajoutant des traductions, des images, etc., plutôt qu'ils ne révisent la/les définition(s), laissant ainsi perdurer des formulations inexactes ou maladroites, ainsi que des erreurs patentées qui pourraient être facilement repérées et corrigées. Face à cette critique, un wiktionnariste actif et administrateur du dictionnaire, présent au colloque parisien d'octobre 2021, plutôt que de remettre en cause les pratiques des contributeurs, tente d'élaborer une ligne de défense : certains contributeurs ne modifieraient pas les définitions « par respect du travail des autres ». Autrement dit, en respectant le travail des autres⁶⁰, le *Wiktionnaire* serait contraire au principe même du wiki⁶¹ sur lequel il repose. On finit par comprendre que le *Wiktionnaire* n'est pas tant le résultat d'un processus d'édition collaboratif, à travers lequel les différents éléments de la microstructure (*e.g.* les définitions) seraient rédigés à plusieurs mains, mais celui d'un processus d'édition agrégatif à travers lequel les articles, et plus globalement le dictionnaire, se construisent par juxtaposition de contributions disparates⁶². En analysant plusieurs articles liés aux mesures d'isolement sanitaire, Sajous & Humbley (2022:185–197) relèvent nombre d'écueils qui tiennent à cette nature plus agrégative que collaborative du processus de rédaction : « dans un même article, chacun apporte sa pierre à l'édifice, sans qu'il y ait de lissage ni de contrôle de cohérence [...] les contributeurs (finalement pas si nombreux) éditent les articles séparément sans que quiconque n'ait de vision globale »⁶³. Sajous et Humbley concluent qu'« il n'y a pas plus de cohérence globale qui émerge du dictionnaire que de cohérence interne aux articles⁶⁴ ». Ainsi, le *Wiktionnaire*, agglomérat d'extraits de sources hétérogènes conçues à différentes époques et de contributions tous azimuts duquel n'émerge souvent aucune cohérence, tient finalement plus d'un patchwork lexicographique que d'un dictionnaire réellement collaboratif.

59. Une comparaison avec l'édition anglaise de *Wiktionary* est prévue.

60. Contrairement, donc, à Landau et ses *second definers* qui se permettent de réviser la formulation du rédacteur initial des définitions.

61. Principe qui permet à tout contributeur d'améliorer un article en ajoutant du contenu, mais aussi – et, peut-être, surtout – en corrigeant le contenu existant.

62. D'où notre préférence pour l'appellation « dictionnaire *contributif* » par rapport à « dictionnaire *collaboratif* » : le *Wiktionnaire* est bien fondé sur les contributions des internautes, mais la supposée collaboration dans le processus rédactionnel est très relative.

63. Caractéristique que le dictionnaire contributif partage avec certains dictionnaires commerciaux et institutionnels.

64. Sont relevés, entre autres, la modification incohérente des marques lexicographiques d'une même série morphologique, l'ajout de sens sans révision (pourtant nécessaire) des sens proches existants et l'ajout d'exemples citationnels qui contredisent la définition qu'ils sont censés illustrer. Ce dernier problème est également constaté par Vincent (2022) dans l'entrée *woke*.

7.2 Diversité : quelle intelligence collective dans les wikis ?

Pour Brabham (2013), comme pour Surowiecki (2005) avant lui, un groupe d'individus peut, dans les bonnes conditions, surpasser aussi bien des experts que les meilleurs individus qui composent ce groupe. Parmi les « bonnes conditions » figurent la diversité du groupe, mais également l'indépendance entre les individus qui composent ce groupe. Nous avons montré (section 5.1) que, dans le *Wiktionnaire*, la diversité est relative. La communauté des contributeurs est par ailleurs un microcosme dont certains membres actifs, parfois géographiquement proches, communiquent en dehors des canaux prévus par l'infrastructure wiki⁶⁵, se rencontrent physiquement, voire collaborent dans le cadre d'autres projets (e.g. le *Dictionnaire des francophones*). L'indépendance des contributeurs est donc, elle aussi, relative.

Une autre condition de succès des projets fondés sur la “sagesse des foules”, ou “l'intelligence collective”, résiderait, selon les auteurs, dans un système adéquat d'agrégation des productions individuelles plutôt que dans la moyenne du travail des individus. Surowiecki (2005) ouvre son livre *The Wisdom of Crowds* par l'étude de l'expérience de Galton (1907), d'où est tirée la citation en exergue. Francis Galton assiste à une foire agricole pendant laquelle un concours consiste à estimer le poids que pèserait un bœuf après qu'il ait été abattu. Galton a l'idée de récupérer, après le concours, les bulletins sur lesquels les participants ont indiqué leur estimation. En calculant la moyenne, il obtient une valeur qui se révèle, à sa surprise (il écrit avoir peu confiance dans le « jugement démocratique »), extrêmement proche du poids réel du bœuf. L'expertise des participants au concours, éleveurs et bouchers habitués à jauger les bêtes, ne suffit pas à expliquer la réussite de l'expérience. D'abord, parce que des visiteurs (non professionnels) de la foire ont également participé au concours. Ensuite, parce que ce type de résultat a été reproduit depuis, notamment à travers l'expérience *jelly-beans-in-the-jar* dans laquelle on demande aux individus d'un groupe d'estimer le nombre de bonbons que contient un bocal. Or, convenons pour cette expérience qu'il est peu probable de trouver des experts en la matière. On peut se demander s'il est réellement pertinent de parler d'intelligence collective pour désigner ce type de « résolution de problème » (*problem solving*). L'intelligence ne revient-elle pas à la personne qui a eu l'idée de calculer la moyenne des estimations pour tenter d'approcher la valeur réelle ? Ces considérations peuvent paraître éloignées des wikis, d'autant que Surowiecki et Brabham s'intéressent majoritairement à d'autres mécanismes que l'édition contributive. La “sagesse des foules”, ou l’“intelligence collective”, sont néanmoins souvent évoquées dans les travaux qui portent sur l'encyclopédie et le dictionnaire contributifs. De plus, un questionnement de Surowiecki (2005) qui intéresse ces ressources est, comme mentionné ci-dessus, la manière dont les contributions individuelles sont mises en commun : moyenne ou agrégation. On parle, pour le *Wiktionnaire* comme pour *Wikipédia*, d'édition collaborative, mais il s'agit en réalité plus d'un processus d'agrégation (cf. section 7.1) dont l'analyse des articles qui en résultent montre les limites. Concernant la moyenne, on ne voit pas en quoi, par exemple, on pourrait “moyenner” les définitions produites par différents contributeurs. Il existe pourtant, comme le relèvent Sajous & Josselin-Leray (2022:351), des initiatives intéressantes, qui permettent l'écriture collaborative de

65. Via notamment Discord, un logiciel de VoIP et de messagerie instantanée. Contrairement aux discussions qui se tiennent dans le cadre du wiki, celles qui utilisent Discord ne sont pas archivées et ne sont donc pas accessibles ultérieurement aux contributeurs absents de la discussion.

définitions. Celle du projet *Kamusi* (Benjamin, 2015), par exemple, repose sur un système de microtâches conçu sous forme d'un jeu qui consiste à gagner des points. Pour un sens à définir, les participants peuvent améliorer une définition existante ou voter pour une définition alternative (ce qui rapporte des points à son auteur). Le mot de la langue cible (autre que l'anglais) est présenté conjointement avec une définition en anglais tirée de *Word-Net*, que les contributeurs peuvent décider de traduire, ou d'ignorer pour proposer une formulation différente. S'il s'agissait de réviser les définitions du *Wiktionnaire*, le « premier jet » existe déjà pour un nombre considérable d'entrées. Et si les contributeurs ont tendance à préférer la nouveauté et délaissier les tâches de maintenance ou d'amélioration qualitative (cf. section 7.3), un système fondé sur la collecte de points, comme proposé dans le projet *Kamusi*, pourrait motiver les contributeurs du *Wiktionnaire*⁶⁶ qui en sont friands : nombreux sont ceux qui collectionnent divers types de badges⁶⁷. À ceux faisant état, par exemple, de l'ancienneté du contributeur, des langues mortes qu'il parle couramment, ou de ses dizaines (voire centaines) de milliers de contributions, pourrait s'ajouter l'indication du nombre de définitions améliorées (l'amélioration étant validée par le vote des autres contributeurs). Par ce moyen ou par un autre, il est nécessaire de persuader les contributeurs de l'importance des définitions et de la nécessité de les réviser. Faute de quoi, si les wiktionnaristes persistent à favoriser uniquement l'ajout d'un vocabulaire en grande partie exotique (hapax, voire mots non attestés, occasionnalismes littéraires, régionalismes désuets, etc.) sans réviser le contenu existant (à commencer par les définitions), le *Wiktionnaire* risque de ne rien devenir de plus qu'un cabinet de curiosités.

7.3 Révis(i)ons : liens entre rigueur, diversité, collaboration et qualité

Nous avons vu en section 3.2 les raisons pour lesquelles plusieurs révisions sont nécessaires à l'écriture des définitions, et pourquoi Landau (2001:382) préconise au moins deux étapes (l'écriture d'un premier jet, puis une révision de la formulation par un autre rédacteur). Les révisions – à plusieurs mains – ne sont pas seulement nécessaires à l'écriture initiale des définitions, ou plus globalement à celle des articles : elles sont également nécessaires à leur mise à jour. Landau (2001:396) compare l'écriture d'un dictionnaire à l'activité de repeindre un pont : dès qu'une couche de peinture est appliquée, le pont en nécessite une nouvelle. Pour Lorentzen & Trap-Jensen (2016), tous les articles devraient être vérifiés au moins de temps en temps, tout élément de l'entrée pouvant faire l'objet de modifications. Mais les wiktionnaristes semblent préférer mener une course à la nomenclature ou ajouter de l'information (pertinente ou non) aux articles que de réviser leur contenu. D'après Lorentzen & Trap-Jensen, cette tendance n'est pas un fait isolé :

new entries and new features attract more attention from users and sponsors alike, because novelties are generally considered more interesting than maintenance, updating is often given less priority and tends to be neglected. This underlines the importance of having a strategy and a tool-box that allow the dictionary staff to diagnose and remedy the parts of the dictionary which are most in need of updating. (2016)

66. Stimuler la motivation des contributeurs est également une des clés du succès des projets fondés sur la participation «des foules».

67. Il s'agit en quelque sorte de décorations pour actes de bravoure wiktionnarienne par lesquels tout contributeur peut faire état de son mérite numérique sur sa « page utilisateur ».

Lorentzen & Trap-Jensen proposent de marquer les entrées du dictionnaire avec plusieurs informations destinées aux lexicographes et aux utilisateurs. Un premier marquage consiste en un repérage des entrées dont la description est susceptible de comporter des éléments obsolètes, afin que ces entrées soient relues à intervalles réguliers. Une case à (dé)cocher, intégrée à l'interface de conception du dictionnaire, vient compléter le dispositif en permettant à l'équipe éditoriale de signaler une entrée qui n'est plus à jour (le membre de l'équipe qui remarque l'obsolescence n'est pas nécessairement en charge de la mise à jour) tout en précisant la nature du problème. Un deuxième marquage, destiné aux utilisateurs, consiste à indiquer les dates de première et dernière révision de chaque entrée. Une mise à jour simultanée de l'intégralité des entrées n'étant pas envisageable, de telles indications permettraient au lecteur de se faire une idée de l'état d'une entrée (à jour/obsolète) et de juger, par exemple, comment interpréter certaines marques (*e.g. now rare, especially younger*) et d'autres éléments en fonction de l'époque de la compilation du dictionnaire (ou, dans le cas de dictionnaires électroniques mis à jour en continu, de celle de l'écriture de chaque article), relativement au moment de consultation.

Même si la microstructure du *Wiktionnaire* n'est pas structurée sous la forme d'une base de données, il serait possible d'y intégrer un mécanisme tel que celui proposé par Lorentzen & Trap-Jensen. Cependant, il n'est pas certain qu'un processus de relecture systématique des articles conduirait au remaniement, souvent nécessaire, des définitions (cf., en section 6.2, la définition problématique de *wagonnet*, inchangée malgré de multiples révisions de l'article par de nombreux contributeurs distincts). Il n'est pas non plus certain que les contributeurs qui estimeraient souhaitable la réécriture d'une définition seraient capables d'en produire une formulation adéquate (cf., en section 6.1, la reformulation par un unique contributeur de la définition d'*islamo-gauchiste*, qui reste problématique, puis la reformulation tout aussi problématique de celle d'*islamo-gauchisme*, par le même contributeur). L'exemple de la contradiction portant sur le genre de *clavardeuse* (cf. section 6.3), en revanche, illustre le cas d'une erreur qui n'aurait pas résisté à un processus de relecture, si celui-ci avait cours dans le *Wiktionnaire*. Ainsi, il y a des chances qu'un nombre insuffisant (ou nul) de révisions soit un indicateur de faible qualité, ce qui donne raison à Landau (2001). En revanche, un nombre important de révisions, même effectuées par une grande diversité de contributeurs, ne garantit aucunement la qualité des articles, ce qui donne en partie tort à Lih (2004), dont les mesures restent néanmoins pertinentes : si elles ne permettent pas de pronostiquer la qualité globale d'un article, la "rigueur" et la diversité sont des indicateurs potentiels de rédaction collaborative, à condition de les mesurer localement, pour chaque élément de la microstructure. Un nouveau questionnement apparaît alors : la collaboration, à ce niveau de granularité, est-elle synonyme de gain qualitatif ? Une étude qui évaluerait la qualité des définitions⁶⁸ les plus révisées (en particulier, celles qui le sont par une grande diversité de contributeurs) permettrait d'examiner la nature de la rédaction collaborative lorsqu'elle porte sur l'activité définitoire, et de vérifier si le principe sur lequel reposent les wikis (la rédaction collaborative) et la diversité potentielle des contributeurs du *Wiktionnaire* ont un impact positif sur la qualité des descriptions lexicographiques.

68. La difficulté d'une telle étude consiste en premier lieu à établir des critères pour évaluer la "qualité" d'une définition (ou d'un article), la littérature étant peu abondante sur ce sujet.

Remerciements

Merci à Pierre Corbin pour sa relecture attentive et ses remarques constructives. Les éventuels problèmes restants trahiraient une prise en compte imparfaite de celles-ci (ou mon étêtement inopportun). Merci également à Nadine Vincent et Chiara Molinari pour l'organisation des très inspirants colloques *Dictionnaires et culture numérique dans l'espace francophone* et pour l'édition de leurs actes.

Les résultats quantitatifs présentés dans cet article ont été produits en utilisant la plateforme OSIRIM administrée par l'IRIT (voir <https://osirim.irit.fr/>).

Références

- Anthony, D., Smith, S. W. & Tim, W. (2005). Explaining Quality in Internet Collective Goods: Zealots and Good Samaritans in the Case of Wikipedia. Fall 2005 Innovation & Entrepreneurship Seminar, MIT.
- Atkins, B. T. S. & Rundell, M. (2008). *The Oxford Guide to Practical Lexicography*. Oxford University Press, Oxford.
- Becker, H. (2015). Scientific and Technical Dictionaries; Coverage of Scientific and Technical Terms in General Dictionaries. In Durkin, P. (Ed.), *The Oxford Handbook of Lexicography*. Oxford University Press, Oxford.
- Béjoint, H. (1998). Scientific and Technical Words in General Dictionaries. *International Journal of Lexicography*, 1(4):354–368.
- Benjamin, M. (2015). Crowdsourcing microdata for cost-effective and reliable lexicography. In *Proceedings of the 9th International ASIALEX Conference*, Hong Kong.
- Boulangier, J.-C. & L'Homme, M.-C. (1991). Les technolectes dans la pratique dictionnaire générale. quelques fragments d'une culture. *Meta*, 36(1):23–40.
- Brabham, D. C. (2013). *Crowdsourcing*. MIT Press, Cambridge.
- Daxenberger, J. & Gurevych, I. (2012). A Corpus-Based Study of Edit Categories in Featured and Non-Featured Wikipedia Articles. In *Proceedings of COLING 2012*, pp. 711–726, Mumbai, India.
- Encyclopædia Britannica (2006). Fatally Flawed: Refuting the Recent Study on Encyclopedic Accuracy by the Journal *Nature*. [https://corporate.britannica.com/britannica_nature_response.pdf].
- Galton, F. (1907). Vox populi. *Nature*, 75:450–451.
- Giles, J. (2005). Internet Encyclopaedias go Head to Head. *Nature*, 438:900–901.
- Hanks, P. (2012). Corpus evidence and electronic lexicography. In Granger, S. & Paquot, M. (Eds.), *Electronic Lexicography*, pp. 57–82. Oxford University Press, Oxford.
- Kittur, A. & Kraut, R. E. (2008). Harnessing the Wisdom of Crowds in Wikipedia: Quality through Coordination. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work, CSCW '08*, pp. 37–46, San Diego, California.
- Kuhn, H. W. (1955). The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2:83–97.
- Landau, S. I. (2001). *Dictionaries: The Art and Craft of Lexicography*. Cambridge University Press, Cambridge.

- Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10(8):707–710.
- Lew, R. (2014). User-generated content (UGC) in online English dictionaries. *OPAL*, 4:8–26.
- Lih, A. (2004). Wikipedia as Participatory Journalism: Reliable Sources? Metrics for evaluating collaborative media as a news resource. In *Proceedings of the 5th International Symposium on Online Journalism*, Austin, Texas.
- Lorentzen, H. & Trap-Jensen, L. (2016). What, When and How? - the Art of Updating an Online Dictionary. In Margalitadze, T. & Meladze, G. (Eds.), *Proceedings of the 17th EURALEX International Congress*, pp. 138–145, Tbilisi, Georgia.
- Mahoudeau, A. (2022). *La Panique woke. Anatomie d'une offensive réactionnaire*. Éditions Textuel, Paris.
- Meyer, C. M. (2013). *Wiktionary: The Metalexigraphic and the Natural Language Processing Perspective*. PhD thesis, Technische Universität Darmstadt.
- Rundell, M. (2017). Dictionaries and crowdsourcing, wikis, and user-generated content. In Hanks, P. & de Schryver, G.-M. (Eds.), *International Handbook of Modern Lexis and Lexicography*. Springer, Berlin, Heidelberg.
- Sajous, F. (2022). Using Wiktionary revision history to uncover lexical innovations related to topical events: Application to Covid-19 neologisms. In Klosa-Kückelhaus, A. & Kernerman, I. (Eds.), *Lexicography of Coronavirus-related Neologisms*, volume 163 of *Lexicographica Series Maior*, pp. 275–306. De Gruyter, Berlin, Boston.
- Sajous, F., Calderone, B. & Hathout, N. (2020). Extraire et encoder l'information lexicale de Wiktionary : quel boulot pour étrangler le goulot! *Lexique*, 27:121–144.
- Sajous, F. & Hathout, N. (2015). GLAWI, a free XML-encoded Machine-Readable Dictionary built from the French Wiktionary. In *Proceedings of the eLex 2015 conference*, pp. 405–426, Herstmonceux, England.
- Sajous, F., Hathout, N. & Josselin-Leray, A. (2018a). Wiktionnaire, un dictionnaire crowdsourcé trop neutre? In *Colloque international Dictionnaires et Culture Numérique dans l'Espace Francophone.*, Milan, Italie.
- Sajous, F., Hathout, N. & Josselin-Leray, A. (2019). *Du vin et devin dans le Wiktionnaire : neutralité de point de vue ou neutralité et point de vue? Études de linguistique appliquée*, 194(2):147–164.
- Sajous, F. & Humbley, J. (2022). Mesures d'isolement sanitaire dans Wiktionnaire et Wikipédia : néologie et lexicographie ou néonymie et terminographie? *Estudios Románicos*, 31:175–201.
- Sajous, F. & Josselin-Leray, A. (2022). Issues in Collaborative and Crowdsourced Lexicography. In Jackson, H. (Ed.), *The Bloomsbury Handbook of Lexicography*, pp. 343–358. Bloomsbury Publishing, London.
- Sajous, F., Josselin-Leray, A. & Hathout, N. (2018b). The Complementarity of Crowdsourced Dictionaries and Professional Dictionaries viewed through the Filter of Neology. *Lexis*, 12.
- Stvilia, B., Twidale, M. B., Smith, L. C. & Gasser, L. (2005). Assessing information quality of a community-based encyclopedia. In *Proceedings of the 2005 International Conference on Information Quality (ICIQ 2005)*, pp. 442–454, Cambridge, Massachusetts.

- Surowiecki, J. (2005). *The Wisdom of Crowds*. Anchor Books.
- Vincent, N. (2021). Quand le lexicographe s'adresse à l'utilisateur. Regard sur l'utilisation des remarques dans deux dictionnaires généraux du 21^e siècle. *La linguistique*, 57(1):133–152.
- Vincent, N. (2022). Faut-il adapter les dictionnaires à l'air du temps? Proposition d'un traitement polyphonique du mot *woke*. Regards linguistiques sur des mots polémiques, *Circula*, 15:122–145.
- Wilkinson, D. M. & Huberman, B. A. (2007). Assessing the value of cooperation in Wikipedia. *First Monday*, 12(4).
- Wolfer, S. & Müller-Spitzer, C. (2016). How Many People Constitute a Crowd and What Do They Do? Quantitative Analyses of Revisions in the English and German Wiktionary Editions. *Lexicos*, 26:347–371.