


Ressources lexicographiques numériques et corpus dans la didactique des langues étrangères : le cas des mots à charge culturelle partagée

 dorif.it/reperes/eleonora-marzi-ressources-lexicographiques-numeriques-et-corpus-dans-la-didactique-des-langues-etrangees-le-cas-des-mots-a-charge-culturelle-partagee/

Eleonora MARZI

Eleonora Marzi

Università di Bologna

Resumé

L'enseignement des langues étrangères fait face aujourd'hui à de nouveaux défis méthodologiques engendrés par une diffusion croissante des ressources linguistiques technologiques et par une approche pédagogique caractérisée par la centralité des compétences pragmatiques dans une perspective interculturelle. Dans ce contexte, le concept de lexiculture formulé par Robert Galisson, qui considère les mots comme des dispositifs d'accès à un réseau culturel implicite et partagé, constitue un cadre de référence sur lequel élaborer des propositions pédagogiques innovantes.

La présente contribution s'inscrit dans cette lignée de recherche et propose une méthode d'enseignement mixte composée par l'usage de ressources lexicographiques numériques (élaborées selon une approche top-down et de nature normative) et de corpus (construits selon une approche bottom-up et interrogeables quantitativement), mettant en avant leur fonctionnalité dans le traitement des mots à charge culturelle partagée. La présentation de deux études de cas – le terme réel «poutine» et l'expression idiomatique « amuser la galerie » – mettent en évidence à la fois les étapes méthodologiques qui les composent et le développement de compétences linguistiques, métalinguistiques, interculturelles et informatiques.

Abstract

The teaching of foreign languages is currently confronted with new methodological challenges arising from the increasing diffusion of technological linguistic resources and from an educational approach characterized by the centrality of pragmatic skills within an

intercultural perspective. In this context, the concept of « lessiculture » formulated by Robert Galisson, which regards words as devices providing access to an implicit and shared cultural network, serves as a framework upon which to develop innovative pedagogical proposals. This contribution aligns with this research direction and proposes a mixed teaching method incorporating the use of digital lexicographic resources (developed through a top-down and normative approach) and corpora (constructed through a bottom-up approach and quantitatively queryable), highlighting their functionality in dealing with words carrying shared cultural significance. The presentation of two case studies – the actual term « poutine » and the idiomatic expression « amuser la galerie » – simultaneously emphasizes the methodological steps that compose them and the development of linguistic, metalinguistic, intercultural, and computer skills.

Introduction : corpora et ressources lexicographiques numériques pour la didactique des langues étrangères

La révolution numérique en cours, englobant notamment la mise à disposition des données linguistiques et des infrastructures pour leur gestion, partage et interrogation, suscite des interrogations cruciales quant aux pratiques d'enseignement des langues étrangères et à l'utilisation des ressources numériques (Grosbois 2012, Molinari 2021). La modification du support intervient dans le processus épistémologique en engendrant le paradigme technocratique qui intègre les technologies dans le processus de conception des activités l'enseignement et l'apprentissage des langues étrangères (Puren 2004).

En fonction du principe sous-tendant la construction des ressources numériques pour la gestion des connaissances linguistiques, nous pouvons identifier deux catégories : d'une part, les corpus, représentant une variété ou un phénomène linguistique et construits selon une approche ascendante (bottom-up), d'autre part, les ressources lexicographiques et terminologiques numériques à vocation normative, élaborées selon une approche descendante (top-down).

La diffusion de l'utilisation des corpus – collections de données représentatives d'un phénomène ou d'une variété linguistique (Biber 1993) – dans l'enseignement des langues étrangères est vaste et diversifiée dans ses applications (Ruggia et al. 2023, Di Vito 2015, Boulton 2007, 2014). Par exemple, les corpus peuvent être utilisés pour étudier le comportement en contexte d'un mot et enrichir le lexique, extraire des termes dans le domaine de l'enseignement des langues spécialisées; les corpus multilingues parallèles peuvent être employés pour des études en traductologie. En plus du renforcement des compétences linguistiques, l'utilisation des corpus influe également sur la méthodologie d'enseignement elle-même, car elle conduit au développement de processus d'apprentissage par induction (Auzéau et al. 2018) ou en autonomie (Ciekanski 2019), sans négliger les compétences informatiques et métalinguistiques (Cavalla 2019).

Notamment, l'utilisation des corpus L2, spécifiquement construits avec des matériaux provenant d'apprenants, joue un rôle crucial dans le suivi de l'apprentissage d'une langue seconde par des locuteurs non natifs. Ces corpus sont souvent annotés en fonction des erreurs et des caractéristiques prosodiques, permettant ainsi, par exemple, d'identifier l'influence de la langue maternelle sur l'apprentissage.

La facilité croissante dans la collecte des données, et la mise à disposition subséquente de corpus via des infrastructures européennes telles que *CLARIN*^[1] ou des consortiums nationaux comme *Ortolang*^[2], qui cataloguent et rendent accessibles une variété de corpus – monolingues, plurilingues, parallèles et comparables, thématiques – ainsi que le développement généralisé d'interfaces d'interrogation (et d'annotation) de plus en plus conviviales, telles que *TXM*^[3], *AntConc*^[4], *LancsBox*^[5], indiquent la voie empruntée caractérisée par l'accès ouvert, l'interopérabilité et la centralisation des accès, tout en préservant leur diversité.

Le format numérique des corpus permet un enrichissement grâce à l'annotation automatique (Schmid, 1994) avec des catégories morphosyntaxiques (lemmes, parties du discours, entités nommées) ou une annotation manuelle avec des catégories sémantiques répondant à des besoins de recherche spécifiques. L'enrichissement représenté par ces annotations est exploité à travers des interrogations rendues possibles grâce au protocole universel CQL – Corpus Query Language (Christ, 1994, 1999), un langage de requête composé d'une syntaxe combinant des paramètres à différents niveaux (morphologique, syntaxique et sémantique) et intégrant des expressions régulières, offrant ainsi une grande flexibilité dans les requêtes.

Les interrogations permettent, par exemple, d'indexer les termes dans le but d'enrichir le lexique, ou d'extraire les concordances d'un terme pivot (Sinclair 1991, Pincemin 2006), afin d'identifier le comportement syntaxique en contexte. Une autre interrogation permet le calcul des collocats (Bertels et al. 2012 ; Mayaffre et al., 2012) qui fournissent des informations sur la prosodie sémantique, une aura de polarité générée par les mêmes collocats d'un terme (Louw, 1993), ou la préférence sémantique liée au champ sémantique généré par les mêmes collocats (Stubbs 1995, Partington, 2004). Dans les deux cas, il s'agit d'intuitions inaccessibles par une *close reading*, d'informations qui ne se trouvent pas dans les ressources lexicographiques, mais qui émergent d'analyses quantitatives et qualitatives de vastes ensembles de données, par le biais de l'extraction de données linguistiques et de la recherche de patrons.

Alors que les corpus contiennent des informations contextuelles et interrogeables sur l'usage des mots en contexte à grande échelle, les ressources lexicographiques numériques fournissent des informations normatives, historiques et sur les variations d'usage (Villers 2006). Leur rôle dans l'enseignement a toujours été central (Nied Curcio 2022), et leur transformation en format numérique, dépassant la simple modification de support, soulève de nouvelles questions (Molinari 2021). L'innovation induite par le changement de support,

du papier au numérique, se manifeste à deux niveaux : le premier concerne l'extension matérielle, permettant, sans la contrainte physique de la page, une inclusion virtuellement infinie d'informations. Le second réside dans une amélioration de l'accès : le format numérique offre une architecture de connaissance conçue pour l'exploration, renforçant ainsi les compétences métalinguistiques. Une expression linguistique peut ainsi être identifiée en tant qu'appartenant à un domaine spécifique, en raison de sa représentation diatopique, diachronique, diastratique ou diaphasique, ou en tant qu'élément d'un réseau associatif ou sémantique issu d'un mot externe à l'expression elle-même. Ces caractéristiques sont largement exploitées dans l'enseignement des langues, où les ressources lexicographiques sont utiles pour obtenir des informations sur le comportement morphosyntaxique d'un mot, le référent qu'il désigne, les variations d'usage possibles, ainsi que la structure du lexique en termes d'hyponymes, d'hyperonymes et de relations de méronymie (Tremblay et al., 2023). De plus, le format numérique permet l'émergence et la diffusion des ressources collaboratives (Murano 2019, Calvi et al. 2023).

Comparées aux corpus, les ressources lexicographiques numériques se caractérisent par des informations détaillées et normatives : un lemme peut être expliqué à travers son étymologie, ses variables d'usage, ou des exemples en contexte choisis pour leur représentativité. Les corpus fournissent une quantité plus importante de données linguistiques – tels que des exemples en contexte ou des collocats qui enrichissent le champ sémantique du lemme –, mais ces informations ne revêtent pas de caractère normatif.

Une telle complémentarité nous conduit à les identifier comme tous deux nécessaires à une approche pédagogique se concentrant sur la gestion de l'aspect culturel contenu dans le lexique, fondamental dans le cadre d'une communication interculturelle mais non explicité de manière transparente dans les ressources linguistiques. La présente contribution propose deux études de cas sur la gestion de mots à charge culturelle partagée, mettant en évidence comment une approche mixte composée de l'utilisation de ressources lexicographiques (top-down) et de corpus (bottom-up) convient de manière complémentaire au développement des compétences linguistiques et métalinguistiques dans une perspective interculturelle. L'article est structuré comme suit : après une introduction à la lexiculurologie et au concept de mots à charge culturelle partagée (§1), suivra l'introduction des études de cas avec la présentation des ressources lexicographiques numériques et des corpus utilisés (§3). Nous nous pencherons ensuite sur le cas du réalisme « poutine » (§4) et de l'expression idiomatique « amuser la galerie » (§5), pour ensuite passer aux conclusions (§6).

1. Les mots à charge culturelle partagée et l'approche mixte

Au début du XXI^e siècle, en plus des innovations technologiques, le domaine des études sur l'enseignement des langues étrangères a vu l'émergence de l'approche actionnelle qui « consiste à considérer les apprenants comme des utilisateurs de la langue et comme des acteurs sociaux, et par conséquent à envisager la langue comme un moyen de communication plutôt que comme un objet d'étude » (Conseil de l'Europe, 2021, p. 29). Bien

que l'importance accordée au contexte et donc à l'aspect pragmatique dans l'enseignement s'était déjà affirmée au milieu du siècle précédent avec l'approche communicative (Puren 2009), l'approche actionnelle adoptée par le Conseil de l'Europe dans le Cadre européen commun de référence pour les langues renforce l'importance du lien indissoluble entre langue et culture, mettant l'accent sur les compétences interculturelles (Demougin 2008, Ciliberti 2012, Rochebois 2017).

Dans ce cadre conceptuel, la notion de lexiculturologie ou pragmatique lexiculturelle, une branche de la didactique des langues postulée par Galisson (1988, 2000, 2002), prend toute sa valeur. Elle se concentre sur le lexique et a pour objet d'étude la lexiculture, soit : « la culture en dépôt dans ou sous certains mots, dits culturels, qu'il convient de repérer, d'explicitier et d'interpréter » (Galisson 2000, p. 52). Les mots sont considérés comme des dispositifs activant des réseaux sémantiques culturels partagés par une communauté et contribuant à un imaginaire collectif, dont l'accès est nécessaire pour décoder des contenus implicites. La culture revêt une dimension spatiale à laquelle on peut accéder :

comme on entre dans la culture par l'ethnologie, l'anthropologie, la sociologie, la sémiologie, l'histoire, la géographie, la philosophie, etc, l'ambition pédagogique est ici d'entrer dans la culture par les mots, afin de solidariser, d'intégrer langue et culture dans un même enseignement apprentissage [...]. (Galisson 2000, p.50).

Les mots, véritables points d'accès, délimitent des « sites lexiculturels », des zones physiques du lexique caractérisées par une forte charge culturelle : bien que tous les mots possèdent une dimension culturelle, certains d'entre eux ont une charge culturelle qui est, d'une certaine manière, « plus lourde, plus évidente, plus disponible que la charge culturelle des autres [mots] » (Galisson, 1991, p. 120). Parmi les particularités des mots à forte charge culturelle, on trouve leur valeur implicite, qui ne se superpose pas au concept de connotation (ayant une dimension individuelle) et ne peut être confondue avec celui de sens (pour lequel les ressources lexicographiques sont adéquates). Les sens latents, parfois dérivés d'autres sens, composent une aura culturelle autour du mot, une aura qui intervient dans l'attribution du sens. Ainsi, les mots permettent un double niveau d'accessibilité :

les mots en tant que réceptacles préconstruits, donc stables et économiques d'emploi par rapport aux énoncés à construire, sont des lieux de pénétration privilégiés pour certains contenus de cultures qui s'y déposent, finissent par y adhérer, et ajoutent ainsi une autre dimension à la dimension ordinaire des signes (Galisson, 1988, p. 331).

Le concept de mot à charge culturelle partagée a récemment été repris par Pamies (2017), qui l'a défini comme culturème, lui attribuant une valeur symbolique extra-culturelle. Agissant tel un modèle métaphorique, le culturème motive l'émergence d'expressions figuratives au sein du langage (Pamies, 2017). La dimension matérielle du lexique tire sa légitimité d'une représentation culturelle qui transcende la langue, mais qui demeure accessible par le biais de celle-ci :

culturemes involve only idiomatic meanings, whose literal referent had previously a symbolic function “outside” language. It is a kind of intersemiotic code shifting (from non-linguistic knowledge towards language). For example, the laurel as a cultural symbol of success and glory necessarily precedes lexical metaphors such as eng. *to rest on one’s laurels* (Pamies 2017, p. 101).

Malgré la centralité et l’importance de la charge culturelle au sein du lexique, les informations qui y sont associées sont extrêmement difficiles à localiser et à gérer. La dispersion des informations culturelles est notable en raison de l’hétérogénéité au niveau de la diffusion – références plus ou moins universelles, au niveau de la stabilité – références culturelles plus ou moins anciennes, ou diastratique – références culturelles propres à différents groupes sociaux.

L’adoption d’une approche mixte par l’utilisation de ressources lexicographiques numériques et de corpus, en raison de leur complémentarité – une approche normative et descendante d’un côté, et une approche d’utilisation réelle et ascendante de l’autre – peut être adaptée pour étudier l’aspect culturel du lexique en extrayant des informations tout en développant simultanément des compétences en informatique, métalinguistiques et interculturelles.

La charge culturelle d’un mot est un ensemble de facteurs qui doit être reconstitué à travers l’identification d’ancrages linguistiques. Il est donc crucial de mobiliser des compétences diverses visant à recueillir et à mettre en relation des données hétérogènes. En détail, il sera nécessaire d’atteindre les points suivants :

- 1) La compréhension du référent culturo-spécifique que le mot désigne ;
- 2) La compréhension des règles morphosyntaxiques du mot ;
- 3) La compréhension des aspects historico-culturels du mot ;
- 4) La compréhension des variétés d’usage ;
- 5) L’étude du comportement de ce mot dans les contextes d’usage ;
- 6) Les éventuelles prosodies sémantiques ou préférences sémantiques ;
- 7) Le positionnement culturel, à considérer non pas comme une simple traduction, mais comme une réintégration dans un univers culturel différent de celui d’origine.

En réponse à ces besoins, les ressources lexicographiques permettent d’extraire des informations nécessaires pour mener des analyses sémiologiques du référent culturo-spécifique que le mot désigne. Elles facilitent également l’extraction d’informations sur les usages normés et les origines étymologiques. Concernant les corpus, ils autorisent l’extraction et l’interprétation de données quantitativement significatives sur le comportement syntaxique des mots, ainsi que sur les prosodies et préférences sémantiques, afin de mieux comprendre leur utilisation dans des contextes non normés. En effet, les corpus enregistrent les variations d’usage de manière plus rapide que les ressources lexicographiques. De plus, à travers les corpus multilingues et parallèles, il est possible d’explorer les phénomènes de traduction et de repositionnement culturel. La diversité typologique des ressources

lexicographiques (dictionnaires extensifs, encyclopédiques, analogiques) et des corpus (généraux/spécialisés, mono/plurilingues) implique une attention particulière lors du choix des ressources à utiliser, qui doit être en adéquation avec les besoins linguistiques. En ce qui concerne le choix des ressources lexicographiques, il est primordial de considérer la diversité pour garantir un niveau d'information aussi complet que possible. Pour les corpus, le principe du choix pertinent prévaut : il est essentiel de connaître les différents types existants afin de sélectionner l'expression du phénomène linguistique la plus adaptée à l'application didactique spécifique. Dans la présentation des deux études de cas suivantes, les ressources choisies, ainsi que les critères utilisés, seront exposés en fonction des cas d'étude spécifiques, tout en détaillant les étapes méthodologiques d'analyse, avec les objectifs et compétences développés par l'apprenant.

2. Les cas d'étude entre realia et expressions idiomatiques

La décision d'utiliser comme études de cas un realia et une expression idiomatique est motivée par les compétences qu'ils peuvent mobiliser et par la forte charge culturelle qu'ils renferment, tout en demeurant des cas connus sur lesquels il est possible de réaliser une contre-épreuve. Les realia, réalités culturo-spécifiques propres à un territoire donné qui contribuent à son identité (Florin 1993), et les expressions idiomatiques, formations linguistiques caractérisées par le principe de non-compositionnalité sémantique (Makowska 2022), résultent de phénomènes culturels solidifiés et représentent d'excellents terrains d'étude pour l'extraction d'informations à contenu culturel. Deux niveaux de compréhension sont nécessaires : le premier concerne la compréhension du référent, et le second concerne l'explication des références culturelles qui se reflètent dans le comportement du mot et dans le contexte d'usage. Nous introduisons par la suite les ressources choisies en fonction des besoins de la recherche.

2.1. Description des ressources

Comme mentionné précédemment, le choix des ressources lexicographiques numériques et des corpus à utiliser est une étape qui doit être réalisée avec la plus grande attention, en liant les exigences didactiques des cas linguistiques spécifiques aux caractéristiques des différentes typologies de ressources.

Les ressources lexicographiques se distinguent par la typologie des informations fournies, chacune répondant à des besoins différents (Lehmann *et al.*, 2018) : des dictionnaires extensifs aux dictionnaires analogiques en passant par les dictionnaires encyclopédiques, spécialisés dans une variété de la langue ou englobant toutes les variétés francophones. Pour commencer, citons le dictionnaire étendu Trésor de la langue française – TLFi^[6], qui permet une navigation – grâce aux balises avec lesquelles il est annoté – selon les variations d'usage et stylistiques. Ensuite, le Petit Robert^[7] dans sa version numérique de 2022 a été utilisé, caractérisé par des renvois analogiques qui contribuent à créer un réseau sémantique associatif bénéfique à l'enrichissement lexical (Zotti, 2014). Mentionnons

également le dictionnaire encyclopédique Larousse^[8], fondamental pour sa contribution à la compréhension du référent extra-linguistique (Kasar, 2008). Pour assurer une couverture étendue de la francophonie, le dictionnaire Usito^[9], dédié à la réalité linguistique québécoise (Cajolet-Laganière *et al.* 2014), et la Base de données lexicographiques panfrancophone^[10], subdivisée en zones géographiques (Amérique du Nord, Maghreb, Caraïbes, Europe, Afrique), ont également été pris en considération.

En ce qui concerne les corpus monolingues, le FrTenTen, sous-corpus de la plus vaste famille de corpus comparables TenTen Family (Jakubíček *et al.*, 2013), a été sélectionné pour la langue française. Ces corpus, collectés à partir du web, sont considérés comme représentatifs d'une typologie linguistique répandue et synchronique (Gatto, 2014). Le corpus est composé d'environ 15 millions de mots et annoté par lemmes et parties du discours. En ce qui concerne le corpus multilingue aligné, OpenSubtitles, un corpus parallèle d'environ 4 millions de mots (Lison *et al.*, 2016), créé à partir de sous-titres de films, a été utilisé. Pour les interrogations le logiciel SketchEngine (Kilgarriff *et al.*, 2004, 2014) a été employé principalement pour deux raisons : son interface user-friendly, très adaptée à des fins didactiques pour effectuer des interrogations, et la possibilité d'accéder à différents corpus consultables en accès ouvert. Par la suite, pour chacun des deux exemples, seront détaillés les activités, les objectifs et les compétences développées.

3. Le *realia*: le cas de la « poutine »

Ce premier exemple est consacré à la gestion du *realia* « poutine » qui est un plat typique de la cuisine québécoise à base de pomme de terre, fromage et sauce brune. La première étape consiste en la recherche au sein des ressources lexicographiques des informations nécessaires d'abord pour comprendre le référent (procédant ainsi à une analyse sémique) et ensuite pour appréhender l'utilisation normalisée du terme.

L'objectif est de développer une compréhension des aspects métalinguistiques de la langue, tels que les types de variation (diachronique, diatopique, diastatique, diaphasique), le concept d'hyponymie, hyperonymie et méronymie, ainsi que les constructions syntaxiques. Les compétences développées visent à croiser et harmoniser des informations provenant de différents niveaux (syntaxique, étymologique, d'usage) et de supports de différentes natures, stimulant ainsi une lecture transversale et intégrative permettant de recomposer dans un cadre conceptuel unique des informations issues de différents aspects.

En ce qui concerne l'analyse sémantique, un support essentiel est fourni par le dictionnaire encyclopédique, le Larousse, où « poutine » est ainsi défini :

poutine = Au Canada, mélange de pommes de terre frites et de fromage en grains arrosé de sauce chaude^[11]

L'analyse sémique permet de déduire que : 1) la poutine est composée de fromage, de frites et d'une sauce chaude, 2) qu'elle est un plat typique d'Amérique du Nord, et plus particulièrement du Canada. La même entrée dans le Petit Robert permet d'apporter deux clarifications: la première est sur le référent – la sauce de plus que chaude est brune – la deuxième est sur la provenance : la Canada a laissé la place au Québec d'où la « poutine » est indiquée comme populaire.

poutine [putin] nom féminin

ÉTYM. 1978 ◇ p.-ê. dérivé de l'ancien français *pou* « bouillie de farine (d'avoine, de maïs) », du latin *puls* « bouillie »

Famille étymologique ⇒ POUDRE.



■ Plat composé de frites garnies de fromage en grains, nappé de sauce brune, populaire au Québec.

Bien que la variation diatopique ne soit pas explicitée, la référence au Québec motive l'utilisation du dictionnaire Usito:

[1](DANS LA CUISINE QUEBÉCOISE) Mets à base de frites garnies de fromage en grains et nappées de sauce brune.

« nous avons englouti des frites, de la poutine et des hot dogs, la joyeuse gastronomie des fins de veillée » (L. Hamelin, 1994).

– *Poutine italienne*, nappée de sauce bolognaise.

[2] Q/C FIG. et FAM. Ensemble d'éléments hétéroclites, d'affaires compliquées, ou d'opérations parfois fastidieuses.

La poutine politique, électorale.

« il a [...] passé trois heures au téléphone avec ses amis du comité de gestion [...] "Il y a beaucoup de poutine interne à régler," a-t-il expliqué » (Le Soleil, 2004).

ÉTYMOLOGIE: Depuis 1978 (*in* DHFQ); depuis 1810 (*in* DHFQ) au sens de « dessert à base de farine ou mie de pain »; origine discutée.^[12]

Le premier sens de « poutine » est similaire à celui du Petit Robert et du Larousse, à l'exception d'une spécificité dans le domaine culinaire et de l'ajout de l'attribut « italienne », faisant référence à une variante de la recette qui contient la « sauce bolognaise », à base de sauce tomate et viande hachée, typique de la ville de bologne que en italien s'appelle « ragù ». Un deuxième sens figuré est présent, où « poutine » désigne un ensemble d'éléments hétérogènes et complexes, avec une connotation négative déduite de l'expression « opérations parfois fastidieuses », caractérisé par une variable distractive « FAM » indiquant une utilisation dans le registre familier. Il n'est que dans ce cas que la variation diatopique

est explicitée par la marque « Q/C », représentant le Québec/Canada, autrement dit: bien que la poutine soit reconnue en dehors du Québec en tant que plat typique, c'est à l'intérieur du Québec que le terme revêt une signification différente, désignant un ensemble hétérogène d'éléments avec une connotation négative, utilisé principalement dans le registre familier. La recherche, menée à travers la Base lexicographique panfrancophone en raison de sa large couverture, compte 15 fiches de définition, dont nous ne reprenons que quelques exemples significatifs.

poutine =

[1] Acadie. Dessert de pâte sucrée et de raisins secs qui cuit dans un sac de coton.

[2] Rare Mélange peu appétissant de divers aliments, des restes le plus souvent.

[3] Fig. Affaire compliquée, organisation complexe; ensemble d'opérations dont la gestion peut être source de difficultés, d'ennuis.

[4] Région. Mets fait de boulettes de pâte farcies de viande hachée et cuites dans un bouillon, une sauce à la viande.

[5] Plat composé de frites et de fromage en grains nappés d'une sauce brune.

L'analyse sémique et la recherche des variations diachroniques, diaphasiques, diastratiques et diatopiques revêtent une importance cruciale. D'un point de vue didactique, elles offrent des informations sur l'usage et les extensions de sens possibles. La corrélation de ces informations suggère deux domaines sémantiques dans lesquels s'inscrit le realia « poutine » : le premier domaine concerne son origine gastronomique avec de nombreuses variations régionales, tandis que le second étend son sens figuré pour désigner un ensemble d'éléments non homogènes. Ce sens figuré peut s'appliquer aussi bien à des objets animés qu'inanimés, et tire son inspiration de l'apparence du plat. Cette extension sémantique est caractéristique des ressources lexicographiques francophones.

La deuxième étape vise à examiner le comportement du terme dans son contexte, avec une attention particulière portée à l'extraction éventuelle de prosodies sémantiques ou d'usages non attestés dans les ressources lexicographiques numériques. Cela vise à enrichir le cadre normatif avec une compréhension du véritable usage en contexte, en intégrant des cas où aucune trace n'est présente dans les ressources lexicographiques numériques peut-être en raison de leur récence ou de leur non-conformité aux normes établies. La réalisation de cette deuxième tâche implique l'interrogation des corpus, en particulier l'extraction des concordances et le calcul des collocats. Cette activité vise à développer une compétence informatique et métalinguistique à travers la formulation de requêtes CQL (Corpus Query Language) et la capacité à interpréter les données linguistiques résultantes. L'extraction des concordances, regroupées ultérieurement en structures syntaxiques, permet d'engager une série de réflexions sur le lexique associé au terme « poutine » (cf. tableau 1).

	Structure syntactique	Fréq.	Index	Concordances
1	à + le	123	2.200	au bacon au homar au boeuf au fois gras au poulet Québec
2	à	47	0.840	Montreal trous
3	de + le	37	0.660	du Québec du monde
4	de	31	0.560	Montréal Banquise Drummondville XX kilos
5	en	25	0.450	en ville en France en compagnie
6	avec	9	0.160	fromage sauce
7	dans	8	0.140	restaurant restos
8	à+les	8	0.140	crevettes bluets
9	chez	5	0.090	Ashton* (nom d'un restaurant)
10	sur	4	0.070	menu menus
11	sans	3	0.050	fromage

Tableau 1 – Concordances de « poutine» regroupées par structures syntactiques (Corpus TenTenFR)

En observant le lexique de chaque structure syntaxique, on constate que la polarité des deux champs sémantiques se divise entre le domaine culinaire et le domaine géographique. Cette constatation est confirmée en recherchant les collocats en filtrant les prédicats où « poutine» est complément d'objet (tableau 2) et sujet (tableau 3).

Tableau 2 – Collocats/ predicats ayant « poutine» comme complément d'objet (Corpus TenTenFR)

Tableau 3 – Collocats/ predicats ayant « poutine» comme sujet (Corpus TenTenFR)

Dans les deux cas, les collocats appartiennent au champ sémantique du domaine culinaire, confirmant ainsi les informations présentes dans les ressources lexicographiques numériques, et enrichissant le lexique corrélié.

La dernière étape vise le repositionnement interculturel du concept par le biais de l'interrogation de corpus multilingues parallèles. Il est important de souligner que, en fonction du type de corpus choisi – celui des sous-titres – l'interrogation ne vise pas la « traduction

correcte» d'autant plus que pour les realia cela est une opération non réalisable (Plassard 2021), mais plutôt le développement d'une compétence interculturelle qui permet de réfléchir sur les stratégies traductives employées pour le repositionnement culturel du concept, et des compétences métalinguistiques liées à ces stratégies. Ce passage est interculturel dans la mesure où replace le concept au sein de la culture cible tout en réfléchissant métalinguistiquement au type de traitement traductif. Nous présentons ci-dessous (tableau 4) quelques exemples explicatifs extrait du corpus aligné français et italien d'OpenSubtitles :

collocate	freq.	index
	985	17.650
râper	32	4.700
déguster	48	3.950
manger	127	3.570
goûter	21	3.220
maudire	18	2.910
danser	10	2.330
savourer	8	1.740
habiller	9	1.690
cuisiner	5	1.630
digérer	3	1.590
dévorer	5	1.390
anti	13	0.960
garnir	4	0.860
concocter	3	0.830
servir	24	0.540
essayer	11	0.420
inventer	11	0.390
enfiler	3	0.220

collocate	freq.	index
	417	7.470
saucer	3	6.470
manger	7	1.400
consommer	3	0.800
vendre	5	0.330
évoquer	8	0.300

	Concordances alignées	Stratégie de traduction
1	a) Ce <i>poutine</i> est délicieux. b) Questo <i>piatto di poutine</i> è delizioso.	traduction approximative - paraphrase
2	a) On a mangé de la <i>poutine</i> , regardé toute la saison, et j'ai pris au moins deux kilos. b) Abbiamo visto tutta la serie mangiando <i>patatine</i> e ho preso, tipo, due chili.	traduction approximative - hyperonymie
3	a) Quelqu'un a mentionné de la <i>poutine italienne</i> , il y a environ une heure. b) Un'ora fa ho sentito qualcuno parlare di <i>patatine col chili</i> .	traduction approximative - hyperonymie
4	a) Les <i>Québécois</i> appellent ça de la <i>poutine</i> . b) I <i>francesi</i> lo chiamano <i>poutine</i> .	traduction approximative - généralisation

Tableau 4 – Concordances parallèles de « poutine » (Corpus OpenSubtitles FR/IT)

Dans le premier exemple [1], la stratégie traductive utilisée est celle de l'approximation par la paraphrase : le nom « poutine» est conservé et est complété par « piatto di», qui permet une contextualisation spécifique dans le domaine culinaire. De même, dans l'exemple [2], la stratégie traductive est celle de l'approximation, mais elle se distingue du cas précédent par l'utilisation de la généralisation, où la « poutine» devient « patatine», l'un de ses composants. Une autre variante est présentée dans l'exemple [3], où la technique d'approximation est réalisée non seulement par la généralisation « patatine», mais aussi par l'ajout d'un parallèle culturel visant à rendre l'effet exotique de l'attribut «italienne» en le remplaçant par « chili». Bien qu'il n'y ait aucune analogie entre les deux, car «italienne» fait référence à la sauce « bolognaise» à base de tomate et viande hachée, tandis que le « chili » est une sauce épicée à base de poivrons, l'aspect intéressant réside dans leur appartenance à une culture autre, élément qui suscite le même effet d'étrangeté auprès d'un public italoophone que « italienne» suscite chez un public francophone. Dans le dernier exemple [4], la stratégie de la traduction approximative est appliquée non à l'objet, mais à la culture, nivelant ainsi l'origine géographique avec l'appartenance linguistique (les Québécois parlent français et sont donc Français), tout en laissant le terme « poutine» inchangé.

Cet exemple illustre comment les ressources lexicographiques numériques et les corpus peuvent être des outils précieux pour comprendre la contribution culturelle d'un mot. Dans le cas spécifique de l'étude, le realia « poutine» se réfère normalement au domaine de la cuisine populaire et traditionnelle, et bien que son utilisation puisse également avoir un sens figuré pour décrire une condition chaotique et parfois négative, dans l'échantillon de langue utilisée, il n'y a pas de trace de cette utilisation. Sans remettre en question les informations lexicographiques, cette dernière observation suggère que le sens figuré n'est pas couramment utilisé.

4. L'expression idiomatique: « amuser la galerie »

Le deuxième exemple se concentre sur l'expression idiomatique « amuser la galerie », qui indique l'acte – mené par un individu – d'entretenir un public. Comme pour le cas précédent, l'objectif est de comprendre le contexte d'usage et l'éventuelle charge culturelle déposée au sein de cette expression.

La particularité des expressions idiomatiques est que, grâce au principe de non-compositionnalité, leur sens global n'est pas le résultat de la somme de sens singuliers des mots. Cependant, souvent, les ressources lexicographiques – selon leur typologie – détaillent le sens des locutions, sous l'une des entrées qui composent les mots.

En observant l'entrée « amuser » du Petit Robert nous trouvons:

1 VX Occuper en faisant perdre le temps.

◆ Tromper en donnant de faux espoirs. « *Les promesses trompeuses dont le faux prophète [...] amusait le peuple* » (Bossuet).

◆ MOD. Détourner l'attention de. → **distraire**. *Tu amuseras le caissier pendant qu'on ouvrira le coffre*. LOC. *Amuser le tapis* : jouer petit jeu en attendant la partie sérieuse. PAR EXT. Donner le change.

2 (XVII^e) Distraire agréablement, donner de l'agrément ; faire rire ou sourire. → **délasser, divertir**. « *Ce jeu de ballon m'a [...] amusé comme un enfant* » (Gide). « *Cela l'amusait beaucoup, il était pris de fou rire* » (Loti). → **égayer**. *Un rien l'amuse*. LOC. FAM. *Amuser la galerie* : faire rire l'assistance en concentrant l'attention sur soi.

▫ **P. p. adj.** *Un sourire amusé*, qui exprime l'amusement.

▫ (Dans une situation où l'on n'est pas d'accord) *Tu m'amuses* (cf. Laisse-moi rire !). *Si ça t'amuse* : si tu veux, si tu en as envie (cf. Si ça te chante).

La première définition, « Occuper en faisant perdre le temps », tend à avoir une connotation négative, tandis que la seconde, « distraire agréablement, donner de l'agrément : faire rire ou sourire », se caractérise par une connotation positive et inclut la définition de l'expression idiomatique : « faire rire l'assistance en concentrant l'attention sur soi », avec la spécification de l'usage familier (« locution familière »). Le TLFi, en tant que dictionnaire extensif, fournit de nombreuses informations et détaille la variété du spectre sémantique, corrélé au comportement syntaxique. Il distingue entre les usages transitifs et pronominaux, les variables diachroniques et diastratiques, ainsi que les emplois techniques, en particulier dans les domaines de l'horticulture, des jeux et du théâtre. L'expression « amuser la galerie » se trouve à l'intérieur du sens qui considère l'objet de l'action comme une personne ou une collectivité, avec la définition de « égayer ».

amuser =

2. Égayer

..... [exemples]

– *Familier*

a) *Amuser la galerie*. Amuser son entourage en concentrant son attention sur soi :

« [...] le premier soir, je le vis tout tranquillement allumer à l'ongle de son petit doigt sa cigarette, puis, comme il venait de perdre au jeu, extraire de mon oreille et de mon nez autant de roubles qu'il fallut, ce qui me terrifia littéralement, mais amusa beaucoup la galerie, car il disait, toujours de ce même air tranquille : « Heureusement que cet enfant est une mine inépuisable! »

A. GIDE, *Les Caves du Vatican*, 1914, p. 740.

Les informations sur l'aspect positif et la marque d'usage familière sont ici confirmées, ainsi que le sens de l'expression idiomatique. Notre recherche s'oriente maintenant vers l'extraction d'informations culturellement liées. À l'intérieur de la fiche « galerie » du Petit Robert, plusieurs sous-domaines sont enregistrés :

galerie =

1. Lieu de passage ou de promenade, couvert, beaucoup plus long que large, ménagé à l'extérieur ou à l'intérieur d'un édifice ou d'une salle. *Galerie autour d'un bâtiment.* → **péristyle**. *Galerie vitrée.* → **véranda**. *Galerie ouverte, cintrée, voûtée, à arcades.* → **portique**. *Les galeries du Palais-Royal.* → **arcade**.
 2. (1893)(Canada) Balcon couvert qui s'étend sur toute la largeur d'une maison. « *une galerie, où les habitants vont se bercer pendant les chaudes soirées estivales* » (M. Laberge).
 3. Salle où sont réunies des collections. *Grande galerie du Louvre*.
- PAR EXT. Magasin où sont exposés des objets d'art en vue de la vente. *Galerie d'art, de peinture* (→ **galeriste**). *Exposer dans une galerie*.
- ◆ PAR MÉTON. Collection d'objets d'art ou de science dans un musée. *Les galeries du Muséum*.
4. (d'abord au jeu de paume) Emplacement réservé aux spectateurs ; les spectateurs eux-mêmes.
- ◆ PAR EXT. Le monde, l'opinion. → **auditoire, public, spectateur, témoin**. *Parler, poser pour la galerie. Il fait cela pour amuser, épater la galerie*.
- ◆ Dans un théâtre, Balcon à encorbellement, à plusieurs rangs de spectateurs. *Premières, secondes galeries*. → **paradis, poulailler**.
- *Galerie d'une église*. → **jubé, tribune**.
5. Cadre métallique fixé sur le toit d'une voiture et qui sert de porte-bagage. *Fixer des skis sur une galerie*.
 6. Passage souterrain ou couvert, pratiqué par l'assiégeant pour s'approcher d'une place. → 2. **sape**.

Du domaine plus général lié à l'architecture en tant que lieu de passage (1, 2) au domaine de l'art pictural et du commerce (3) ainsi qu'au sport (4). À l'intérieur de ce dernier, l'expression « amuser la galerie » est insérée sous la définition qui attribue par extension à la galerie le sens de « public » ou « auditoire ». Cependant, ce sens est lié au type de public : dans le cas d'un jeu – le jeu de paume, ancêtre du tennis – et dans le cas du théâtre. Les renvois analogiques montrent comment au théâtre les galeries étaient familièrement appelées « paradis » ou « poulailler », faisant référence à leur hauteur par rapport à la scène

et le bruit qui s'en dégagait. Les galeries étaient les rangées les plus élevées, destinées au peuple et qui en raison de la faible visibilité étaient à des prix très bas, contrairement aux loges où les plus aisés avaient droit à des fauteuils ou des sièges.

poulailler =

2 (1803) FAM. Galerie supérieure d'un théâtre. → **paradis**. « *les habitués de l'amphithéâtre suprême, vulgairement dit poulailler* » (Nerval). *Prendre une place au poulailler.*

paradis =

5 (1606) Galerie supérieure d'un théâtre. → **poulailler**. « *Les Enfants du Paradis* »

En poursuivant la recherche dans le dictionnaire encyclopédique Larousse, nous trouvons confirmation que la « galerie » désignait la partie la plus haute du théâtre, à une époque ancienne:

galerie = (théâtre) Partie la plus voisine des combles, dans les anciens théâtres.

Synonymes : balcon – corbeille – mezzanine – paradis (familier) – poulailler (familier)

Étage situé au-dessus des loges et des fauteuils de balcon.

La collecte d'informations à travers le TFLi confirme les intuitions extraites lors de la consultation des ressources lexicographiques.

B. – 1. HIST. Allée longue et couverte qui bordait l'un des côtés des terrains des jeux de paume et où se tenaient les spectateurs

P. ext. et fam. Ensemble indéterminé de personnes considérées comme spectateurs.

Synon. *monde, public. Étonner, impressionner la galerie; faire des effets (de qqc.)*

devant la galerie. Des loustics d'atelier (...) faisaient sourire la galerie en disant un mot plaisant sur la grimace de chaque cadavre (ZOLA, T. Raquin, 1867, p. 84). *Il amusait la galerie de ses gasconnades* (A. DAUDET, N. Roumestan, 1881, p. 116). *Dans ce métier où l'éloquence occupe la galerie, mais où l'entregent fait le reste, il n'y a pas de petites tâches* (H. BAZIN, Qui j'ose aimer, 1956, p. 109).

2. En partic.

a) Balcon en encorbellement, à plusieurs rangs de spectateurs, sur le pourtour d'une salle de théâtre. *Première, deuxième, troisième galerie. En haut (...) à la troisième galerie, autour de la rotonde du plafond (...) des appels et des rires sortaient d'un brouhaha continu de voix* (ZOLA, *Nana*, 1880, p. 1095). *Vous donneriez cent sous pour trouver un col de chemise, du dixième rang de fauteuils à la deuxième galerie!* (COLETTE, *Vagab.*, 1910, p. 8).

La « galerie » est associée au jeu de paume, et par extension – et dans un langage familier – elle désigne ce type de public. Dans le domaine théâtral, on parle de spectateurs qui se trouvent dans des places qui ne sont pas exceptionnelles en termes de visibilité de la scène, avec des exemples caractérisés par un niveau d'excentricité qui veut être mis en évidence :

« des appels et des rires sortaient d'un brouhaha continu de voix ». En cherchant donc à uniformiser les informations recueillies, nous pouvons affirmer que l'extension du sens de « galerie » avec « public » a des racines historiques, et provient d'un type de public 1) pour le sport et 2) théâtral (en particulier populaire), donc caractérisés par le divertissement. La variation diastratique « familier » est indiquée comme appartenant aux lexies « poulailler » et « paradis », mais aussi au terme « galerie » dans l'acception qui désigne les spectateurs. Par conséquent, la marque d'usage « familier » s'applique à l'ensemble de l'expression. Toutefois cette marque d'usage n'est pas la seule suggérée par les informations extraites ; l'hypothèse est que celui qui « amuse la galerie » le fait dans un contexte joyeux, peu formel, non régi par des manières préétablies ou des étiquettes strictes. Cela parce que le public auquel la « galerie » fait référence est typique du sport et du théâtre (dans sa variété non solennelle) : cette charge culturelle partagée se retrouve à l'intérieur de l'expression idiomatique et pourrait rendre compte d'emplois en contexte qui impliquent une farce, un ton comique ou satirique.

Cette intuition peut être confirmée à travers l'interrogation des corpus, d'abord monolingues, puis multilingues.

Nous avons d'abord calculé les collocats de l'expression entière, obtenant le tableau suivant (tableau 5), et observons comment le champ sémantique de la farce et de la comédie est prédominant.

Collocats	Freq. Cooccurrences	Freq. Candidats	Index LogDice
pitre	20	5 200	6,18
pitreries	9	4 467	5,15
BOUFFONS	4	71	5,07
gesticulation	4	3 527	4,15
clowns	14	25 044	3,99
bouffons	5	7 643	3,84
facéties	4	6 739	3,63
gogos	4	8 856	3,37
épater	5	13 273	3,26
guignol	4	10 101	3,23
imitations	4	14 382	2,85
passe-passe	3	12 906	2,55
distraindre	7	41 110	2,35
divertir	7	41 967	2,32
ridiculiser	3	16 396	2,28
détourner	14	107 038	2,05

Tableau 5 – Collocats pour la locution « amuser la galerie » (Corpus TenTenFR 2018)

Les noms « bouffon », « clown », « pitre » sont des indices de comédie burlesque, et le nom « guignol » fait référence à un programme télévisé satirique de marionnettes français, né dans les années 80, qui parodiait le monde politique, les médias, les personnalités ou plus généralement la société française. Le nom « gogo », indiquant un homme crédule et niais, qui

se laisse tromper facilement, renforce l'idée d'un contexte burlesque. Des termes comme « facéties », « gesticulation », « imitation » renforcent le champ sémantique de la comédie burlesque, « passe-passe » est un tour de prestidigitation utilisé par les prestidigitateurs, les illusionnistes, et au sens figuré, une illusion, une tromperie. Les collocations prédicats confirment le même champ sémantique : « divertir », « ridiculiser », associés à d'autres termes connotés de manière moins forte : « détourner », « distraire ». La présence de « épater » indique une équivalence synonymique avec la locution « épater la galerie ».

La dernière étape de l'interrogation des corpus se fait par l'extraction des exemples du corpus multilingue parallèle et sur leur interprétation. Il est important de rappeler que les résultats des concordances dans les corpus parallèles ne doivent pas être considérés comme normatifs, mais comme des utilisations possibles en contexte, en gardant toujours à l'esprit le type de corpus examiné qui, comme nous l'avons déjà mentionné, est la représentation d'un phénomène linguistique donné. L'activité consiste à faire défiler et à considérer les occurrences significatives pour engendrer des pistes de réflexion. En voici quelques exemples.

	Corcordances alignées	Champ sémantique
1	a) Tu aurais bu trop de punch et tu aurais amusé la galerie. b) Hai bevuto troppo punch e sei stata al centro della festa.	excès
2	a) pour amuser la galerie avec son gros nez rouge. b) che faceva divertire il pubblico col suo grosso naso rosso.	comicité burlesque
3	a) Pete me prend pour un clown bon à amuser la galerie . b) Pete mi considerava buono solo a fare il buffone.	comicité burlesque
4	Avec tout mon respect, sénatrice, cette équipe n'est pas là pour amuser la galerie . Con tutto il rispetto, Senatore, ma questa unità non scherza	politique
5	a) Il nous faut un candidat qui amuse la galerie. b) Ci serve un candidato che metta in piedi uno spettacolo.	politique
6	a) Bien, s'il veut amuser la galerie, jouer à cache-cache avec le sous-marin, sans parler de ce qu'il veut faire avec le mari de la Secrétaire, nous allons faire aussi le spectacle. b) Se vuole pavoneggiarsi, giocare a nascondino con il sottomarino, per non parlare di ciò che sta facendo con il marito del Segretario, dovremmo dare spettacolo anche noi.	politique

Tableau 6 – Concordances parallèles pour l'expression «amuser la galerie» (Corpus OpenSubtitles FR/IT)

Le premier exemple est caractérisé par un contexte d'excès. On peut observer que l'indication de la substance alcoolique « punch », associée à la perte d'inhibitions. Le deuxième champ sémantique identifié est celui de la comicité typique du cirque, avec la référence au nez rouge et au clown. Il semble que l'utilisation en contexte témoigne d'une tendance à l'aspect excessif de la comédie ou du divertissement. Un autre aspect qui se

révèle caractérisant est celui du champ sémantique politique : les contextes des exemples (4), (5), (6) montrent que le lexique politique est bien présent. Les termes « candidato, Sénatrice, Segretario » renvoient à des contextes politiques, reprenant le lien avec le collocat « Guignoles », relative à l'émission télévisée satirique politique.

En résumé, l'expression idiomatique « amuser la galerie » est définie par les ressources lexicographiques comme l'acte d'entretenir un public par un individu. En croisant les informations contenues dans les ressources lexicographiques numériques, nous avons perçu la charge culturelle de l'expression, principalement liée au terme « galerie » qui porte en lui une histoire liée à des publics joyeux, ludiques et décontractés. La confirmation de cette charge culturelle partagée réside dans l'identification de la prosodie et de la préférence sémantique évidentes, qui se manifestent dans l'utilisation de l'expression dans des contextes toujours comiques, allant du cirque à la satire politique.

4. Conclusions

Nous avons décrit une méthodologie à travers deux études de cas utilisant des ressources lexicographiques numériques et des corpus comme outil d'enseignement pour la gestion des mots à charge culturelle partagée. La méthode présentée repose sur l'utilisation d'une approche mixte composée par ressources lexicographiques numériques et corpus.

L'utilisation de l'approche mixte permet de développer des compétences de consultation et d'interprétation de données hétérogènes, acquérir des notions sur le comportement d'un mot au niveau normatif et d'usage à la fois, enrichir simultanément le champ lexical par des associations ou des délimitations du champ sémantique.

De plus, l'utilisation de l'approche mixte renforce l'apprentissage de notions métalinguistiques : savoir reconnaître, par exemple, les variables d'usage, le comportement syntaxique et morphologique, l'étymologie d'un mot, et savoir utiliser ces concepts pour des interrogations dans un environnement technologique. Cela entraîne la mémorisation des notions apprises, et l'apprenant sera amené à se familiariser avec la variété des ressources lexicographiques numériques et des programmes d'interrogation de corpus. Il est très important d'avoir des choix pertinents ; il pourra apprendre à manipuler des CQLs utiles pour interroger toutes sortes de corpus.

La structure en étapes, la spécification des objectifs et des compétences développées, ainsi que les pistes de réflexion permettent de rendre la méthodologie scalable, donc son application à d'autres phénomènes linguistiques. En détail, nos cas d'étude ont montré comment les deux ressources permettent l'accès à des informations diverses et complémentaires dans une optique de pragmatique culturelle. Les mots, considérés comme points d'accès à la culture, peuvent être observés à travers des interrogations précises, tant dans leur dimension normative que dans leur utilisation réelle, dans leur dimension historique et dans leur forme en réseau.

L'examen des exemples que nous avons pris en considération montre comment, à côté d'un sens rapporté comme locution dans les ressources lexicographiques numériques, il existe des informations supplémentaires mais implicites qui trouvent leurs racines directement dans l'imagerie culturelle.

Ces aspects nous semblent valables pour argumenter que l'utilisation d'une approche mixte corpus-lexique peut être efficace et encore riche de dimensions à explorer.

Bibliographie

AUZEAU Fanny, ABIAD, Laure, « Le corpus : un outil inductif pour l'enseignement-apprentissage de la grammaire », *Synergies France* n. 12, 2018, p. 175-187.

BERTELS, Ann, SPLEELMAN, Dirk, « La contribution des cooccurrences de deuxième ordre à l'analyse sémantique », *Corpus* n.11, 2012, DOI : <https://doi.org/10.4000/corpus.2184>.

BIBER, Douglas, « Representativeness in Corpus Design », *Literary and Linguistic Computing*, n. 8, 4, 1993, p. 243 – 257, <https://doi.org/10.1093/lc/8.4.243>.

BOULTON, Alex, TYNE, Henry, *Des Documents authentiques aux corpus : démarches pour l'apprentissage des langues*. Paris, Didier, 2014.

BOULTON, Alex, « Esprit de corpus: Promouvoir l'exploitation de corpus en apprentissage des langues », *Texte et corpus*, n. 3, 2007, p.37- 46.

CAJOLET-LAGANIERE, Hélène, D'AMICO Serge, « Usito : pour un apprentissage enrichi du lexique en contexte québécois », *Québec français*, n. 171, 2014, p. 72-73.

CALVI, Silvia, DANKOVA, Klara, « Lexicographie traditionnelle-collaborative et FLE : compétences lexicales et lexicographiques », n. 86, 2023, DOI : <https://doi.org/10.4000/linx.10108>.

CAVALLA, Cristelle, « Comment former les étudiants de Master FLE à l'utilisation pédagogique des corpus numériques ? », in GOES, J. et al., *Apports et limites des corpus numériques en analyse de discours et didactique des langues de spécialité*, Editura Universitaria, 2019, p.79-92.

CHRIST, Oliver, « A modular and flexible architecture for an integrated corpus query system », *COMPLEX'94*, Budapest, 1994.

CHRIST, Oliver et al., *The IMS Corpus Workbench: Corpus Query Processor (CQP)*, Institute for Natural Language Processing, University of Stuttgart, 1999.

CILIBERTI, Anna, « La Nozione Di 'Competenza' Nella Pedagogia Linguistica: Dalla 'Competenza Linguistica' Alla 'Competenza Comunicativa Interculturale', *Italiano Linguadue*, n. 2, 4, 2012, p 2-10.

CIEKANSKI Maud, « Les corpus : de nouvelles perspectives pour l'apprentissage des langues en autonomie ? », *Recherches en didactique des langues et des cultures*, n. 11, 1, 2014, DOI : 10.4000/rdlc.1710.

CONSEIL DE L'EUROPE, *Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer*, Strasbourg, Éditions du Conseil de l'Europe, 2021.

DI VITO, Sonia, « L'utilisation des corpus dans l'analyse linguistique et dans l'apprentissage du FLE », *Linx*, n. 68-69, DOI : <https://doi.org/10.4000/linx.1519>.

FLORIN, Sider, « Realia in translation », ZLATEVA, Palma (dir.), *Translation as Social Action*, New York, Routledge, p. 122-128, 1993.

DEMOUGIN, Françoise, « La didactique des langues – cultures à la croisée des méthodes », *Tréma*, 2008, DOI : 10.4000/trema.427.

PLASSARD, Freddie « Les realia au cœur du débat traductologique », in *Al-Zaum*, Malek et al. (dir.), A propos des realia. Littérature, traduction et didactique des langues, Editions des Archives contemporaines, 2021.

GALISSON, Robert, « Didactologie : de l'éducation aux langues-cultures à l'éducation par les langues-cultures », *Éla. Études de linguistique appliquée*, n. 128, 4, 2002, p. 497-510, DOI : 10.3917/ela.128.0497.

GALISSON, Robert, « Cultures et lexicultures. Pour une approche dictionnaire de la culture partagée », *Annexes des Cahiers de linguistique hispanique médiévale – Hommage à Bernard Pottier*, n. 7, 1988, p. 325-341, DOI: <https://doi.org/10.3406/cehm.1988.2133>.

GALISSON, Robert, « La pragmatique lexicoculturelle pour accéder autrement, à une autre culture, par un autre lexique », *Melanges Crapel*, n. 25, p. 47-73, 2000.

GATTO, Maristella, *Web As Corpus, Theory and Practice*, London, Bloomsbury, 2014.

GROSBOIS, Muriel, *Didactique des langues et technologies*, Paris, Presses de l'Université Paris-Sorbonne, 2012 ;

JAKUBICEK, Miloš *et al.*, « The TenTen Corpus Family », *7th International Corpus Linguistics Conference*, 2013.

KASAR Sündüz Öztürk, « Pour une typologie moderne des dictionnaires », *Synergies Turquie* n. 1, 2008, p. 55-64.

KILGARRIFF, Adam *et al.*, « The Sketch Engine: ten years », *Lexicography*, n. 1, p. 7-36, 2014.

KILGARRIFF, Adam *et al.*, « The Sketch Engine », *Proceedings of the 11th EURALEX International Congress*, p.105-116, 2004.

LEHMANN, Alise, MARTIN-BERTHET, Françoise, « Chapitre 10. Les types de dictionnaires », *Lexicologie. Sémantique, morphologie et lexicographie*, LEHMANN, Alise, MARTIN-BERTHET, Françoise (dir.), Paris, Armand Colin, 2018, p. 257-266.

LISON, Pierre, TIEDEMANN Jörg, « OpenSubtitles2016: Extracting Large Parallel Corpora from Movie and TV Subtitles », *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC-2016)*, 2016.

LOUW, Bill, « Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosodies », in BAKER, Mona (dir.), *Text and Technology: In Honour of John Sinclair*, Philadelphia/Amsterdam, John Benjamins, p. 157-176, 1993.

MAYAFFRE, Damon, VIPREY, Jean-Marie, (dir.), *La cooccurrence, du fait statistique au fait textuel*, *Revue CORPUS*, n. 11, 2012.

MOLINARI, Chiara, « La lexicographie numérique : un nouvel outil pour l'enseignement du FLE ? », *ELA. Études de linguistique appliquée*, n. 201, 1, 2021, pp. 27-47, DOI : [10.3917/ela.201.0028](https://doi.org/10.3917/ela.201.0028).

MURANO, Michela, « Les dictionnaires collaboratifs et la phraséologie : de la description du patrimoine existant à l'invention de nouvelles séquences figées », *ELA. Études de linguistique appliquée*, n. 194, 2019, p. 193-209. DOI : [10.3917/ela.194.0193](https://doi.org/10.3917/ela.194.0193)

NIED CURCIO, Martina, *L'uso del dizionario nell'insegnamento delle lingue straniere*, Roma, RomaTre Press, 2022.

PAMIES, Antonio, « The Concept of Cultureme from a Lexicographical Point of View », *Open Linguistics*, n. 3, p. 100 – 114, 2017, DOI [10.1515/opli-2017-0006](https://doi.org/10.1515/opli-2017-0006).

PARTINGTON, Alan, « “Utterly content in each other’s company”; Semantic Prosody and Semantic Preference », *International Journal of Corpus Linguistics*, n. 1, 9, Philadelphia/Amsterdam, John Benjamins, 2004, p. 131-156.

PINCEMIN, Bénédicte, « Concordances et concordanciers : de l'art du bon KWAC », XVIIe colloque d'Albi Langages et signification – Corpus en Lettres et Sciences sociales : des documents numériques à l'interprétation, 2006, Albi, France. pp.33-42.

PUREN, Christian, « Entre l'approche communicative et la perspective actionnelle, quoi de neuf ? », in : ABDELGABER, Sylvie, MÉDIONI, Maria-Alice (dir.), *Les Cahiers Pédagogiques, Collection des HNS*, 2009, p. 87-91.

PUREN, Christian, « Quels modèles didactiques pour la conception de dispositifs d'enseignement/apprentissage en environnement numérique ? », *Éla. Études de linguistique appliquée*, 2004/2 (n° 134), p. 235-249. DOI : 10.3917/ela.134.0235.

ROCHEBOIS, Christianne Benatti, « Langue, Culture Et Didactique Du FLE », *Caligrama: Revista de Estudos Românicos*, n. 21, 2, 2017, p. 95.

RUGGIA, Simona. GAILLAT T, Thomas (dir.), « Les corpus numériques pour la didactique des langues : de la formation des enseignants à l'élaboration de dispositifs d'apprentissage », *Revue Corpus*, n. 24, 2023, DOI: <https://doi.org/10.4000/corpus.7438>.

SCHMID, Helmut, « Probabilistic Part-of-Speech Tagging Using Decision Trees », *Proceedings of International Conference on New Methods in Language Processing*, University of Manchester, Manchester, 1994.

SINCLAIR, John, *Corpus, concordance, collocation*, Oxford, Oxford University Press, 1991.

STUBBS, Michael, « Collocations and Semantic Profiles: On the Cause of the Trouble with Quantitative Studies », *Functions of Language*, n. 1, 2, Philadelphia/Amsterdam, John Benjamins, p. 23-55, 1995.

TREMBLAY, Ophélie, FRASSI, Paolo, (dir.), *Cahiers de lexicologie, Dictionnaires, ressources lexicales et didactique des langues*, n. 122, 1, 2023, DOI : <https://doi.org/10.4000/linx.1519>.

VILLERS, Marie-Éva de, « La pratique lexicographique », *Profession lexicographe*, Montréal, Presses de l'Université de Montréal, 2006.

ZOTTI, Valeria, « Les renvois analogiques du Petit Robert : un système sémiotique complexe », *Les sémiotiques du dictionnaire*, Berlin, Frank&Timme, 2014.

[1] <https://www.clarin.eu/resource-families>

[2] <https://www.ortolang.fr/market/corpora>

[3] <https://txm.gitpages.huma-num.fr/textometrie/Documentation/>

[4] <https://www.laurenceanthony.net/software/antconc/>

[5] <http://corpora.lancs.ac.uk/lancsbox/>

[6] <http://atilf.atilf.fr/>

[7] <https://dictionnaire.lerobert.com/>

[8] <https://www.larousse.fr/>

[9] <https://usito.usherbrooke.ca/>

[10] <https://www.bdlp.org/recherche?bases%5B%5D=QU>

[11] Larousse: <https://www.larousse.fr/dictionnaires/francais/poutine/63197> (dernière consultation 1/12/2023).

[12] <https://usito.usherbrooke.ca/d%C3%A9finitions/poutine> (dernière consultation 1/12/2023).

Per citare questo articolo:

Eleonora MARZI, « Ressources lexicographiques numériques et corpus dans la didactique des langues étrangères : le cas des mots à charge culturelle partagée », *Repères DoRiF*, numéro hors-série *Varia*, DoRiF Università, Roma, febbraio 2024, <https://www.dorif.it/reperes/eleonora-marzi-ressources-lexicographiques-numeriques-et-corpus-dans-la-didactique-des-langues-etrangees-le-cas-des-mots-a-charge-culturelle-partagee/>

ISSN 2281-3020

Quest'opera è distribuita con Licenza Creative Commons Attribuzione – Non commerciale – Non opere derivate 3.0 Italia.

