

ACADEMIA ROMÂNĂ

INSTITUTUL DE FILOLOGIE ROMÂNĂ „ALEXANDRU PHILIPPIDE”, IAȘI
INSTITUTUL DE LINGVISTICĂ „IORGU IORDAN – ALEXANDRU ROSETTI”, BUCUREȘTI
INSTITUTUL DE LINGVISTICĂ ȘI ISTORIE LITERARĂ „SEXTIL PUȘCARIU”, CLUJ-NAPOCA

**LEXICOGRAFIA ACADEMICĂ ROMÂNEASCĂ.
PROVOCĂRILE INFORMATIZĂRII
AcadLexi 2025**

Volumul cuprinde lucrări prezentate la ediția a XIV-a a Colocviului cu participare internațională „Lexicografia academică românească. Provocările informatizării”. AcadLexi 2025 (<https://philippide.ro/acadlexi.html>).

Coordonatori de volum: Elena Isabelle Tamba, Mariana Nastasia

Editori:

Alexandru Anghelina
Iulia Barbu-Comaromi
Călina Ciocoiu
Alexandra Ilie
Mircea Minică
Andreea-Nora Pușcaș

Referenți științifici:

Conf. univ. dr. Ana-Maria Minuț
CS I dr. Gabriela Haja

Coperta: Lucian Săcrieru-Dragomir

ISSN: 3008-6159

ISSN-L: 3008-6159

DOI: https://doi.org/10.52257/30086159_2025

© 2025 Editorii volumului. Toate drepturile rezervate.

Universitatea Babeș-Bolyai

Presa Universitară Clujeană

Str. Hasdeu nr. 51

400371 Cluj-Napoca, România

Tel./fax: (+40)-264-597.401

E-mail: editura@ubbcluj.ro

<http://www.editura.ubbcluj.ro/>



ACADEMIA ROMÂNĂ

INSTITUTUL DE FILOLOGIE ROMÂNĂ „ALEXANDRU PHILIPPIDE”, IAȘI
INSTITUTUL DE LINGVICĂ „IORGU IORDAN – ALEXANDRU ROSETTI”, BUCUREȘTI
INSTITUTUL DE LINGVICĂ ȘI ISTORIE LITERARĂ „SEXTIL PUȘCARIU”, CLUJ-NAPOCA

ACTELE COLOCVIULUI INTERNAȚIONAL „LEXICOGRAFIA ACADEMICĂ ROMÂNEASCĂ. PROVOCĂRILE INFORMATIZĂRII” AcadLexi 2025

PRESA UNIVERSITARĂ CLUJEANĂ

2025

LANGUE, CULTURE ET MÉMOIRE : POUR UNE LEXICOGRAPHIE NUMÉRIQUE INTERPRÉTATIVE ET INCLUSIVE

ALEKSANDRA BANJEVIC*

1. Introduction

À l'heure où la numérisation des savoirs transforme en profondeur les pratiques scientifiques, la lexicographie se trouve à un tournant décisif : entre l'efficacité technique offerte par les outils numériques et la nécessité de préserver la complexité du réel linguistique. Ce basculement soulève des interrogations d'une portée épistémologique majeure : comment décrire les langues sans les figer ? Comment intégrer la variation sans l'assujettir aux normes dominantes ?

Nous proposons ici une réflexion critique sur les relations entre lexicographie et mémoire culturelle, en prenant pour terrain d'observation les variétés dialectales du Monténégro. Dans cet espace plurilingue et historiquement stratifié, les pratiques lexicales révèlent une richesse souvent occultée par les dispositifs normatifs traditionnels. À travers l'analyse des emprunts romans et leur intégration dans le tissu linguistique local, ce travail entend explorer la portée culturelle, symbolique et identitaire du lexique.

Loin d'être une entreprise neutre, la lexicographie engage des choix – lexicaux, idéologiques, mémoriels – qui orientent la perception du langage et des communautés qui le portent. Comme l'a souligné Alain Rey, « la lexicographie n'est jamais une entreprise innocente » : elle institue des hiérarchies, entérine des omissions, naturalise des représentations.

Dans cette perspective, nous défendons l'idée d'une lexicographie interprétative, attentive aux dynamiques langagières minorées, et consciente de sa responsabilité dans la préservation du patrimoine immatériel. Les formes dialectales, longtemps marginalisées dans les ressources normatives, deviennent dès lors des objets d'étude à part entière, porteurs d'une mémoire culturelle dont la lexicographie numérique peut, et doit, se faire l'écho.

* Faculté de Philologie, Université „Crne Gore”, Danila Bojovića, 81400, Nikšić, Monténégro.

2. La langue comme matrice mémorielle et culturelle

La langue ne saurait être réduite à une simple fonction instrumentale : elle constitue une archive vivante, un espace de cristallisation des représentations collectives et des expériences historiques. Chaque unité lexicale, chaque tournure idiomatique porte en elle les traces d'une vision du monde propre à une communauté donnée. Cette conception s'inscrit dans le sillage des travaux de Wilhelm von Humboldt et Leo Weisgerber, pour qui la langue structure non seulement la pensée, mais également l'expérience vécue du réel (von Humboldt 1836 ; Weisgerber 1953).

Dans cette optique, la lexicographie acquiert une dimension herméneutique : elle devient un outil d'accès aux strates profondes de la mémoire culturelle. Elle ne se limite pas à désigner les objets ou à nommer les réalités ; elle révèle les configurations symboliques et les héritages perceptifs qui sous-tendent l'usage lexical.

Les théories de la mémoire culturelle développées par Aleida et Jan Assmann offrent ici un cadre pertinent (Assmann, Assmann 2010). Le langage, en tant que vecteur de transmission intergénérationnelle, joue un rôle central dans la constitution des identités collectives. Dans les contextes plurilingues et historiquement mouvants tels que les Balkans, cette fonction mémorielle se manifeste avec une intensité particulière.

Les variétés dialectales du Monténégro illustrent cette dimension mémorielle par leur richesse lexicale stratifiée. En particulier dans les zones frontalières, littorales ou anciennement urbaines, on observe une présence marquée d'emprunts romans – issus de l'italien, du vénitien ou du dalmate – qui ont été absorbés, transformés et réinterprétés par les locuteurs. Ces mots, loin de constituer de simples insertions exogènes, forment des couches sédimentées du lexique local, analogues à des strates archéologiques. Chacun d'eux témoigne d'un contact prolongé, d'une coexistence linguistique et d'une mémoire partagée¹.

3. La lexicographie comme interprétation

Contrairement à la conception répandue d'une activité purement descriptive, la lexicographie s'inscrit dans un processus interprétatif complexe. Définir un mot, choisir un exemple d'usage, hiérarchiser les acceptions : autant de gestes éditoriaux qui impliquent des arbitrages culturels, idéologiques et scientifiques (Béjoint 1994). Le lexicographe n'est jamais un simple observateur ; il opère à l'intersection entre langue vivante et norme écrite, entre usage spontané et codification institutionnelle.

Chaque dictionnaire, qu'il soit imprimé ou numérique, reflète une vision du monde. Les choix d'inclusion ou d'exclusion, la terminologie mobilisée, la structure des définitions – tout cela construit une certaine image du langage, socialement située (Nuessel 1992). Cette perspective rejoint la critique sociolinguistique formulée par Pierre Bourdieu, pour qui la langue est aussi un capital symbolique dont l'accès n'est ni neutre ni égalitaire (Bourdieu 1982).

¹ V. notamment les études dialectologiques sur la côte adriatique monténégrine dans Lopičić Volarić 2016.

Dans les contextes plurilingues et historiquement multicouches comme ceux des Balkans, cette dimension interprétative prend une acuité particulière. Documenter – ou ignorer – certaines formes dialectales ou minoritaires revient à valider ou à marginaliser des identités culturelles. La lexicographie devient dès lors un lieu de négociation entre reconnaissance et invisibilisation, entre inclusion lexicale et effacement normatif (Fillmore, Atkins 2003).

Le choix des sources, en particulier, révèle ces dynamiques. Un corpus centré sur les textes administratifs ou juridiques ne pourra rendre compte de la vitalité lexicale des récits oraux, des usages populaires ou des formes diastratiques. Il est donc essentiel d'adopter une approche plurielle, où la diversité des registres, des géographies et des pratiques est intégrée dans la pratique lexicographique (Hanks 2013).

4. La révolution numérique en lexicographie

L'introduction des outils numériques dans le champ lexicographique a marqué un tournant méthodologique et épistémologique majeur. L'accès à des corpus textuels massifs, la visualisation automatisée des cooccurrences, ou encore l'extraction assistée de structures syntaxiques ont profondément modifié les conditions de description lexicale (Kilgarriff *et al.* 2014). Des plateformes comme le *Trésor de la langue française informatisé* (TLFi), COROLA ou Sketch Engine permettent aujourd'hui une observation fine et extensible de l'usage réel des mots, dans des contextes multiples.

Pendant, cette puissance technologique s'accompagne de défis critiques. Le processus de numérisation n'est jamais neutre : le choix des textes, des genres discursifs et des variétés linguistiques intégrées dans les corpus conditionne les résultats obtenus (Sinclair 1991). Trop souvent, les formes dialectales, orales ou non standardisées sont sous-représentées, voire absentes, des bases de données linguistiques dominantes, ce qui entraîne une reproduction implicite des hiérarchies linguistiques existantes (McEnery, Hardie 2012).

L'uniformisation des données, exigée par les contraintes des environnements de traitement automatique, tend à invisibiliser les variations. Ce phénomène rejoint les analyses de Sarah G. Thomason, selon lesquelles les dynamiques de contact linguistique impliquent également des rapports de pouvoir (Thomason 2001) – désormais transférés dans les architectures algorithmiques elles-mêmes.

En outre, l'automatisation partielle de certaines tâches – telles que la génération de définitions ou l'alignement sémantique – délègue à des modèles computationnels des fonctions traditionnellement assumées par l'expertise humaine. Or, ces modèles sont eux-mêmes construits à partir de corpus préalablement biaisés, ce qui soulève des questions fondamentales d'ordre épistémologique et éthique.

Il devient dès lors impératif de concevoir la lexicographie numérique non comme une fin en soi, mais comme un espace critique où l'outil technique doit rester subordonné à une vision interprétative, contextuelle et inclusive du langage. La technologie doit être au service de la diversité langagière, et non de sa simplification.

5. Étude de cas : les emprunts romans dans les variétés dialectales du Monténégro

Les variétés dialectales du Monténégro, notamment dans les régions côtières, constituent un espace d'observation privilégié pour l'étude du contact linguistique. L'héritage roman y est particulièrement visible à travers un ensemble d'unités lexicales empruntées à l'italien, au vénitien et, dans une moindre mesure, au dalmate, intégrées durablement dans l'usage local (Lopičić Volarić 2016).

Ces emprunts ne se limitent pas à une simple juxtaposition exogène : ils sont morphologiquement adaptés, sémantiquement redéployés et culturellement réinvestis. Leur présence, souvent absente des dictionnaires normatifs, mérite une attention particulière dans une perspective de lexicographie inclusive.

5.1 Tableau d'analyse lexicographique : microstructure des emprunts

ponistra

- étymologie : vénitien *fenestra*
- sens local : fenetre
- integration morpho-phonetique : chute de la syllabe initiale /fe-/, passage de [f] a [p], insertion euphonique
- usage/contexte : usage courant dans le sud du montenegro
- remarques : souvent absente des dictionnaires standards ; forme diastratique regionale

škura

- étymologie : italien *scuro*
- sens local : volet, obscurite
- integration morpho-phonetique : adaptation phonetique complete, lexicalisation d'un sens architectural
- usage/contexte : maisons traditionnelles en pierre
- remarques : glissement semantique a partir du champ visuel vers l'objet architectural

šerveta

- étymologie : italien *servietta*
- sens local : napperon, linge de table
- integration morpho-phonetique : transformation de [v] en [r], forme euphemisee
- usage/contexte : vocabulaire domestique, souvent oral
- remarques : non standardisee, absente des ressources officielles

batalja

- étymologie : italien *battaglia*
- sens local : querelle, dispute verbale

- integration morpho-phonétique : conservation morphologique, specialisation pragmatique

- usage/contexte : discours familial ou de voisinage
- remarques : figement d'un sens figuré, charge expressive populaire

kampanela

- étymologie : italien campanella
- sens local : petite cloche (eglise ou maison)
- integration morpho-phonétique : reduction des consonnes doubles, déplacement de l'accent

- usage/contexte : contexte religieux ou domestique
- remarques : forte charge sonore et symbolique, ancrage mémoriel

5.2 Données de corpus et validation numérique

L'exploitation de Sketch Engine à partir d'un sous-corpus régional (monténégrin littoral) a permis d'identifier la fréquence et les contextes d'usage de certaines unités. Ainsi, *pitura* (de *pittura*) apparaît régulièrement dans des textes liés à la rénovation domestique, l'artisanat ou la décoration intérieure². Bien que ces formes soient rares dans les corpus nationaux standardisés, leur vitalité dans les corpus oraux ou spécialisés témoigne d'un usage enraciné.

Le contraste entre leur présence effective dans les pratiques langagières et leur absence dans les dictionnaires institutionnels illustre l'écart entre lexique vivant et lexique reconnu. Une lexicographie contemporaine soucieuse d'inclusivité devrait, en ce sens, intégrer ces formes dans ses bases, accompagnées d'indicateurs diastratiques et culturels.

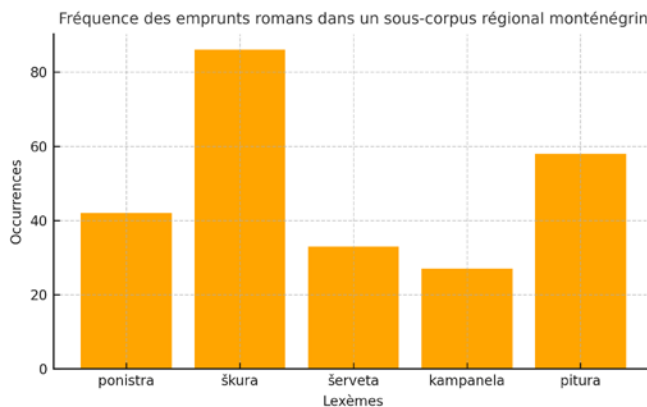


Figure 1. Fréquence des emprunts romans dans le corpus régional (Sketch Engine, 2024).

² Données extraites du sous-corpus monténégrin dans Sketch Engine (requête : *pitura*, 2024), co-occurrences les plus fréquentes : *boja*, *kuća*, *radionica*.

Le graphique ci-dessous illustre la fréquence d'occurrence de cinq emprunts romans dans un sous-corpus régional monténégrin, tel qu'extrait à l'aide de Sketch Engine. Ces données, bien que représentatives de manière indicative, soulignent la vitalité de certaines unités lexicales dans l'usage local.

6. Discussion : pour une lexicographie interprétative et inclusive

Les analyses précédentes ont mis en lumière que les emprunts romans dans les variétés dialectales monténégrines ne sauraient être considérés comme des éléments lexicaux marginaux ou anecdotiques. Leur intégration morphologique, leur réinterprétation sémantique, ainsi que leur persistance dans l'usage quotidien témoignent d'une richesse lexicale durable – à la fois culturelle, historique et identitaire (Kristiansen *et al.* 2008).

Pourtant, ces unités lexicales demeurent largement invisibles dans les dictionnaires normatifs et les corpus de référence. Cette omission révèle les limites d'un modèle descriptif traditionnel, centré sur la norme écrite et les usages majoritaires. Elle invite à reconfigurer les fondements de la lexicographie contemporaine, en particulier dans les contextes plurilingues et dialectalement diversifiés (Svensén 2009).

Nous plaçons ici pour une lexicographie interprétative, entendue comme une pratique qui ne se contente pas de décrire ce qui est, mais qui cherche à révéler ce qui a été effacé, minoré ou relégué. Il ne s'agit pas de produire un dictionnaire dialectal séparé – souvent perçu comme périphérique – mais d'intégrer, au sein même des dictionnaires généraux, des marqueurs diatopiques, diastratiques et socioculturels (Lebart, Salem 1994).

Un modèle inclusif supposerait notamment :

- l'identification et l'étiquetage systématique des formes dialectales attestées dans les corpus spécialisés,
- la documentation de leur variation formelle et sémantique à travers les contextes d'usage,
- la mise en relation entre la forme standard et ses variantes régionales, sans hiérarchie implicite,
- l'intégration de commentaires culturels ou pragmatiques permettant de situer le mot dans son écosystème discursif (Rundell 2003).

Une telle approche, déjà esquissée dans certains projets collaboratifs de lexicographie numérique, rejoint les appels récents en faveur d'une lexicographie « participative », ouverte à la contribution des locuteurs et porteuse de pluralité épistémique.

En donnant place à ces voix minorées, la lexicographie assume un rôle mémoriel et symbolique fondamental. Elle devient un espace de médiation entre langue, mémoire et reconnaissance culturelle – un outil scientifique, certes, mais aussi un geste éthique.

7. Conclusion : vers une lexicographie attentive à la mémoire lexicale

La diversité linguistique ne constitue pas un résidu folklorique du passé, mais une dimension vivante et structurante du patrimoine immatériel des communautés. À travers le

prisme des emprunts romans intégrés dans les parlers monténégrins, cette étude a cherché à démontrer que le lexique agit comme un dépôt de mémoire culturelle, un indicateur d'interactions historiques et un révélateur de dynamiques identitaires.

Loin d'un exercice purement descriptif, la lexicographie se trouve investie d'une fonction interprétative, éthique et inclusive. Elle ne saurait se limiter à l'inventaire normatif des formes légitimes : elle doit aussi rendre visible la richesse des usages périphériques, marginaux ou minoritaires – sans les hiérarchiser ni les folkloriser.

Dans cette perspective, la lexicographie numérique représente une opportunité majeure, à condition qu'elle intègre des principes de contextualisation, de transparence méthodologique et d'ouverture aux variations. En concevant le dictionnaire non pas comme une fin normative, mais comme un espace dialogique entre mémoire et usage, entre tradition et innovation, on lui restitue sa portée culturelle pleine.

Ainsi, dans les contextes linguistiquement stratifiés tels que les Balkans, la lexicographie ne peut être dissociée de la reconnaissance culturelle. Documenter les strates invisibilisées du lexique, c'est contribuer à une mémoire collective plus juste – et à une science du langage plus solidaire.

BIBLIOGRAPHIE

- Assmann, Assmann 2010 = Jan Assmann and Aleida Assmann, *Communicative and Cultural Memory. In A Companion to Cultural Memory Studies*, Berlin, de Gruyter.
- Béjoint 1994 = Henri Béjoint, *Tradition et innovation dans les dictionnaires monolingues*, Paris, Duculot.
- Bourdieu 1982 = Pierre Bourdieu, *Ce que parler veut dire : L'économie des échanges linguistiques*, Paris, Fayard.
- CoRoLa = *Corpusul de referință pentru limba română contemporană*, <https://corola.racai.ro/>.
- Fillmore, Atkins 2003 = Charles J. Fillmore, B. T. Sue Atkins, *Describing frame semantics for lexicography*, in „International Journal of Lexicography”, 16(3), p. 281–296.
- Hanks 2013 = Patrick Hanks, *Lexical Analysis: Norms and Exploitation*, Cambridge, MIT Press.
- Humboldt 1836 = Wilhelm von Humboldt, *Sur la diversité de la structure du langage humain*.
- Kilgariff *et al.* 2014 = Adam Kilgariff, Vít Baisa, Jan Bušta, Miloš Jakubiček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, Vít Suchomel, *The Sketch Engine: Ten Years On*, in „Lexicography”, vol. 1, p. 7–36, <https://doi.org/10.1007/s40607-014-0009-9>.
- Kristiansen, Dirven 2008 = Gitte Kristiansen, René Dirven, *Cognitive Sociolinguistics: Language Variation, Cultural Models, Social Systems*, Berlin, de Gruyter.
- Lebart, Salem 1994 = Luc Lebart, André Salem, *Statistique textuelle*, Paris, Dunod.
- Lopičić Volarić 2016 = Mirjana Lopičić Volarić, *Lingvistički kontakti na crnogorskom primorju*, Podgorica, Académie des sciences et des arts du Monténégro (CANU).
- McEnery, Hardie 2012 = Tony McEnery and Andrew Hardie, *Corpus Linguistics: Method, Theory and Practice*, Cambridge, Cambridge University Press.
- Nuessel 1992 = Frank H. Nuessel, *The Study of Dictionaries and Lexicography*, Westport, Greenwood Press.

- Rundell 2003 = Michael Rundell, *Recent Trends in English Pedagogical Lexicography*, in „International Journal of Lexicography”, no. 16(3), p. 315–342, <https://doi.org/10.1093/ijl/11.4.315>.
- Sinclair 1991 = John Sinclair, *Corpus, Concordance, Collocation*, Oxford, Oxford University Press.
- Sketch Engine = *Sketch Engine: Create and search a text corpus*, <https://www.sketchengine.eu/>.
- Svensén 2009 = Bo. Svensén, *A Handbook of Lexicography: The Theory and Practice of Dictionary-Making*, Cambridge, Cambridge University Press.
- Thomason 2001 = Sarah G. Thomason, *Language Contact: An Introduction*, Washington, Georgetown University Press.
- TLFi = *Trésor de la langue Française informatisé*, ATILF – CNRS & Université de Lorraine, <http://www.atilf.fr/tlfi>.
- Weisgerber 1953 = Leo Weisgerber, *Die Stellung der Muttersprache im Leben des Menschen*.

LANGUAGE, CULTURE AND MEMORY: TOWARDS AN INTERPRETATIVE AND INCLUSIVE DIGITAL LEXICOGRAPHY

ABSTRACT

This paper explores digital lexicography not merely as a tool for linguistic documentation, but as a cultural, memorial, and ethical act. By bringing together etymological, sociolinguistic, and semiotic perspectives, it examines the role of words as collective archives within the dialectal varieties of Montenegro—particularly through the lens of Romance borrowings. Drawing on digital corpora (TLFi, COROLA, Sketch Engine), the study advocates for a lexicographic approach that is sensitive to linguistic diversity, cultural memory, and identity dynamics.

Engaging with the theoretical contributions of Humboldt and Weisgerber, the paper underscores that language is not merely a vehicle for communication but a dynamic space of world-construction. In this light, lexicography becomes an interpretative act, situated at the intersection of scientific inquiry and symbolic meaning – a gesture aimed at preserving intangible heritage and recognizing marginalized voices.

Keywords: *digital lexicography, linguistic diversity, cultural memory, Montenegrin dialectology, Romance borrowings, ethics of language.*

