

GEOLINGVISTICA ROMÂNEASCĂ ÎN ERA DIGITALĂ*

SILVIU-IOAN BEJINARIU^{1,2}

VASILE APOPEI²

FLORIN IFTENE¹

1. INTRODUCERE

Cartografia lingvistică românească are o tradiție de peste un secol, fiind una dintre cele mai bogate din Europa. De-a lungul timpului, tehnologia de realizare a hărților lingvistice a evoluat de la hărți realizate cu tehnologii offset până la cele realizate cu tehnologii informatice publicate atât ca volum tipărit precum și on-line pe internet. În această lucrare este prezentată evoluția tehnicilor folosite pentru publicarea atlaselor lingvistice în România.

Cel mai mare avantaj al utilizării tehnologiilor informative în editarea asistată a planșelor lingvistice nu este neapărat o calitate grafică superioară sau facilitarea editării, cât faptul că stocarea informației lingvistice în baze de date specializate oferă posibilitatea folosirii acesteia în studii sau analize computerizate. Un exemplu în acest sens constă în faptul că indexurile de cuvinte și forme sunt generate în mod automat. De asemenea, trecerea de la un mod de vizualizare la altul (hartă lingvistică – hartă sintetică / interpretativă – material (ne) cartografiat) este realizată automat, fără a fi necesară transcrierea repetată a răspunsurilor din anchetele dialectale. Deoarece poziționarea geografică este una dintre trăsăturile importante ale informației în dialectologie utilizarea aplicațiilor de tip Sistem de Informații Geografice devine o necesitate. Acestea oferă și alte avantaje: eliminarea redundanțelor din informațiile stocate în bazele de date, calitate ridicată a hărților indiferent de scara la care se face afișarea, calitate obținută prin utilizarea descrierilor vectoriale a trăsăturilor geografice, posibilitatea afișării cu diferite niveluri de detaliere precum și posibilitatea de a selecta trăsăturile ce sunt sau nu vizualizate.

În continuare, sunt trecute în revistă principalele realizări din geolingvistica românească de la începuturi și până în zilele noastre, cu un accent pe cele bazate pe utilizarea instrumentelor informatice. Începute în urmă cu 20 de ani, colaborările dintre dialectologi și informaticieni au permis nu doar crearea de instrumente de cercetare care au ușurat munca dialectologilor, dar și identificarea unor noi direcții de cercetare și colaborare. În acest sens menționăm colaborările rodnice pe care Institutul de Informatică Teoretică al Academiei Române, Filiala Iași, le are cu departamentul de Dialectologie, geografie lingvistică, fonetică și onomastică de la Institutul de Lingvistică al Academiei Române „Iorgu Iordan – Al. Rosetti”, coordonat de prof. univ. dr. Nicolae Saramandu, membru corespondent al Academiei Române.

* “This work presented in this paper was partially supported by the research grant from the Romanian National Authority for Scientific Research and Innovation CNCS = UEFISCDI, project no PN-III-P4-ID-PCE-2020-0451”

2. EVOLUȚIA CARTOGRAFIEI LINGVISTICE ÎN ROMÂNIA

2.1. Începuturile cartografiei lingvistice în România

Preocupări legate de studiul variației limbii au existat cu mult înainte de publicarea primului atlas lingvistic. Încă din 1716, Dimitrie Cantemir face în *Descriptio Moldaviae* o analiză „Despre graiul moldovenilor”. Pornind de la studiul legăturii dintre graiurile moldovenesc, latinesc și italianesc, Cantemir remarcă și diferențele în rostire care depind de gradul de cultură, sex dar și localizarea geografică și învecinarea cu vorbitori ai altor limbi. B. P. Hașdeu a realizat începând cu 1884 prima anchetă dialectală prin corespondență în scopul elaborării unui Dicționar Enciclopedic al Limbii Române (*Etymologicum Magnum Romaniae*), operă neterminată din care a publicat patru volume ajungând doar până la definiția cuvântului *bărbat*. Iar acestea nu sunt singurele exemple.

Primul atlas lingvistic al dialectului dacoromân a fost publicat de romanistul german Gustav Weigand între 1898 și 1909 sub formă de nouă fascicule: *Linguistischer Atlas des dacorumänischen Sprachgebietes* (WLAD). După cum este subliniat și în (Petrovici 1954: 5), aceasta a fost prima abordare sistematică în studiul dialectului dacoromân. Weigand stabilește și ariile dialectale corespunzătoare (sub) dialectelor bănățean, muntean și moldovean. Se remarcă faptul că pentru Ardeal nu este identificat un dialect propriu, autorul considerând că nu există zone dialectale unitare de mare întindere, cu excepția Maramureșului. Cele 67 de hărți WLAD au fost realizate pentru un chestionar de 130 cuvinte și o rețea de 752 puncte. În 48 dintre hărți sunt analizate formele fonetice pentru cele 6 sectoare în care a fost împărțit teritoriul studiat. Un exemplu de hartă pentru zona III (de nord) este prezentat în Figura 1.a. Variația diatopică a unor fenomene fonetice este ilustrată în 16 hărți. De asemenea, sunt incluse hărți ce prezintă repartitia dialectală (Figura 1.b), traseele anchetelor și componența etnică (Olariu F.-T., Olariu V. 2010: 90).

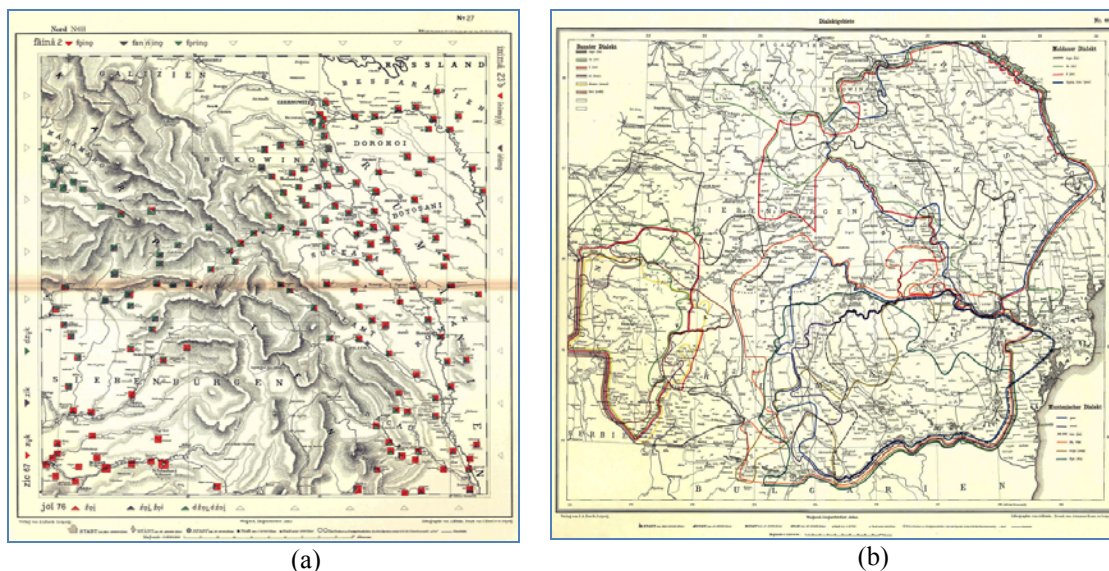


Figura 1. Hărți WLAD: (a) hartă cu forme fonetice în dialectul moldovenesc, (b) hartă cu izoglose pentru trei dialecte reprezentative. (Sursa: (a) harta 27 în WLAD: 46, (b) harta 65 în WLAD: 84, Public Domain)

Transcrierile fonetice sunt prezentate pe laturile hărții însoțite de simboluri de formă triunghiulară care sunt plasate și pe harta lingvistică pentru localizarea în punctele de anchetă.

de un milion de fișe, autorii au intenționat publicarea atlasului în zece volume. Din cauza celui de-al doilea război mondial, din cele zece volume nu au apărut decât trei: două volume din *ALR I* în 1938 și 1942, și unul singur din *ALR II* în 1940. Pentru *ALR II* au fost publicate și două suplimente din care unul conținea texte dialectale în transcriere fonetică.

Un element de noutate în *ALR* este includerea între punctele de anchetă a unui număr de 15 localități cu limbi minoritare, ceea ce a fost de interes pentru studiul altor limbi și dialecte din sud-estul Europei. Din punctul de vedere al realizării grafice, se remarcă faptul că formatul stabilit în *ALR* pentru planșele lingvistice s-a păstrat până astăzi cu foarte puține modificări. Informațiile suplimentare sunt organizate în trei *Note* care conțin întrebarea din cursul anchetei (*Nota I*), completări ale răspunsurilor obținute de la informatori (*Nota II*) și alte explicații și completări ale autorului / anchetatorului (*Nota III*), la fel ca în atlasele lingvistice actuale. Fiecare dintre cele trei volume *ALR* cu hărți analitice (Figura 2) a fost însoțit de câte un volum denumit *Micul Atlas Lingvistic Român* (ALRM) care conține hărți sintetice color (Figura 3). Un alt element de noutate introdus de *ALR* este reprezentat de publicarea de planșe cu material ne-cartografiat.

2.3. Atlasele lingvistice românești în ultima jumătate de secol

Materialul rămas nepublicat din *ALR I* și *ALR II* a fost valorificat sub conducerea lui Emil Petrovici de Institutul de Lingvistică din Cluj-Napoca. Astfel, între 1956 și 1972 au fost publicate șapte volume ale *Atlasului Lingvistic Român – serie nouă*. De asemenea, au fost publicate și patru volume ale *Micului Atlas Lingvistic Român – serie nouă*, până în 1981. Trebuie menționat că aceste volume conțin doar materialele colectate de Emil Petrovici pentru *ALR II*, cele colectate de Sever Pop pentru *ALR I* nefiind încă publicate.

În 1957 Emil Petrovici propune realizarea *Noului Atlas Lingvistic Român pe Regiuni* (NALRR) având ca model *Le Nouvel Atlas Linguistique de la France, par régions* (NALF). Pentru a crea o imagine fidelă a structurii dialectale, rețeaua de puncte a fost îmbogățită ajungându-se la aproximativ 1000 de puncte de anchetă, dar numărul de întrebări din chestionar a fost redus comparativ cu cel din *ALR*, la 2600 incluzând aici și întrebările din chestionarul introductiv (Olariu F.-T., Olariu V. 2010: 96). *NALRR* conține un număr de șapte atlase regionale pentru dialectul dacoromân și un atlas pentru dialectele sud-dunărene. Anchetele s-au desfășurat pe parcursul a aproximativ 20 de ani (1960-1980, cu tot cu anchetele pe probă) cu participarea cercetătorilor din București, Cluj-Napoca, Iași și Timișoara. În paralel, din 1967 a început și publicarea volumelor *NALRR*. Titlul acestora a variat în timp în funcție de preferințele autorilor și editorilor dar și ca rezultat al cenzurii din perioada regimului politic comunist: *Noul atlas lingvistic român, pe regiuni* (NALRR), *Noul Atlas lingvistic român* (NALR), *Atlasul lingvistic român, pe regiuni* (ALRR), *Noul Atlas lingvistic al României* (Olariu F.-T., Olariu V. 2010: 97). Astfel, au fost publicate 5 volume din *NALR*. *Oltenia*, 4 volume din *ALRR*. *Maramureș*, 4 volume din *NALR*. *Banat*, 4 volume din *NALR*. *Moldova și Bucovina*, 4 volume din *ALRR*. *Transilvania*, 5 volume din *ALRR*. *Muntenia și Dobrogea*, 4 volume din *NALR*. *Crișana*, ca să le amintim doar pe cele dedicate dialectului dacoromân și care au fost publicate în România. Autorii acestora au încercat să aducă diverse îmbunătățiri ale modului de prezentare și sistematizare a materialului lingvistic, cele mai multe dintre ele fiind apoi preluate în volumele publicate ulterior: hărți sintetice și interpretative fonetice, morfologice sau lexicale cu arii delimitate de izoglose ce conțin simboluri și hașuri, hărți lingvistice tradiționale cu transcrieri fonetice combinate cu izoglose, diferite moduri de

grupare și ordonare a răspunsurilor în planșe cu material (ne) cartografiat dacă nu există un interes deosebit în ceea ce privește repartitia geografică (Saramandu 2007: 96). După știința noastră, pentru dialectul dacoromân mai sunt în lucru sau în curs de publicare volume din *NALR – Moldova și Bucovina și ALRR – Muntenia și Dobrogea*.

Nu în ultimul rând, trebuie menționată și seria de atlase lingvistice publicate de cercetătorii dialectologi din Republica Moldova. *Atlasul Lingvistic Moldovenesc (ALM)* publicat între 1968 și 1987 include două volume (fiecare cu câte două părți), 6 volume de texte dialectale și un dicționar dialectal în 5 volume. Deși *ALM* a fost publicat folosind alfabetul chirilic, el este apropiat de seria *ALR* prin metodologia utilizată la elaborare. După 1990, cu un nou titlu – *Atlasul lingvistic român, pe regiuni. Basarabia, Nordul Bucovinei, Transnistria* și editat folosind alfabetul latin (Figura 4), atlasul publicat de Institutul de Lingvistică al Academiei de Științe a Moldovei „face parte din seria acestor atlase regionale” (Pavel 1993: 163).

Alte lucrări deja publicate și dedicate celorlalte dialecte ale Limbii Române sunt: *Atlasul lingvistic al graiurilor românești din Banatul iugoslav* de Radu Flora, *Micul atlas lingvistic al graiurilor istroromâne* de Radu Flora, *Micul atlas al dialectului aromân din Albania și din Fosta Republică Iugoslavă Macedonia* de Petru Neiescu, *Atlasul lingvistic al dialectului meglenoromân* de Petar Atanasov. Este în curs de elaborare și *Atlasul lingvistic român. Graiurile dintre Morava, Dunăre și Timoc* la Institutul de Lingvistică și Istorie Literară „Sextil Pușcariu” din Cluj-Napoca.

Atlasul lingvistic al dialectului aromân (ALAR) de Nicolae Saramandu și Manuela Nevaci (editor) este de asemenea o lucrare realizată pe baza chestionarului *NALR* cu 2543 de întrebări, „făcând posibilă compararea materialului dialectal aromân cu cel cuprins în atlasele regionale” (ALAR: V). Primul volum a fost publicat în anul 2014, iar cel de-al doilea volum a apărut în anul 2020.

Având în vedere faptul că unele fenomene studiate în atlasele regionale capătă o semnificație mai accentuată prin evidențierea ariilor dialectale la nivel național Nicolae Saramandu pune încă din 1990 „problema elaborării unui atlas al atlaselor lingvistice regionale sau a unei sinteze a Noului atlas lingvistic român pe regiuni (*NALR. Sintează*)” (Saramandu 1990: 57(1)). Primul volum a fost publicat în 1998: *Atlasul lingvistic român pe regiuni* –

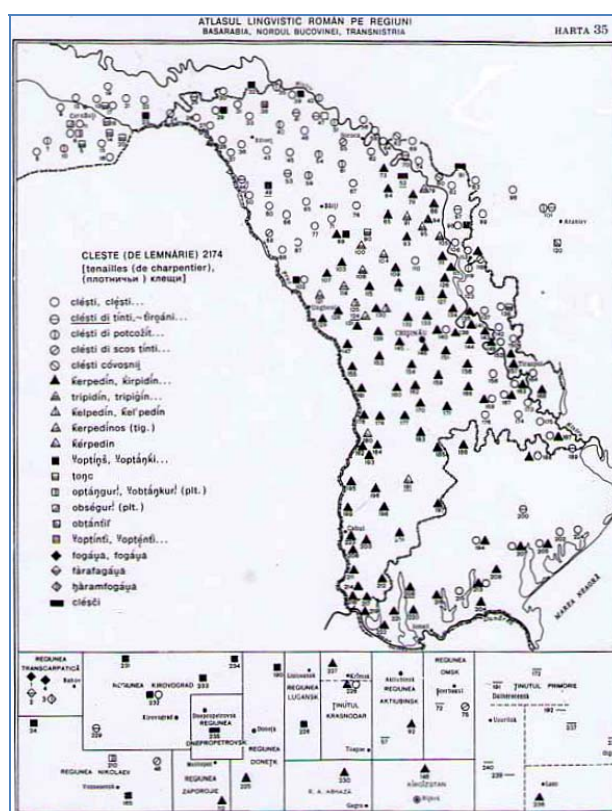


Fig. 4. Hartă din *Atlasul lingvistic român pe regiuni. Basarabia, Nordul Bucovinei, Transnistria*. (Sursa: Pavel, 2012: 35)

Sinteză (ALRR. Sinteză), coordonat fiind de Nicolae Saramandu. Rețeaua este formată din aproximativ 1200 puncte, inclusiv cele din R. Moldova, iar hărțile sunt interpretative.

Majoritatea covârșitoare a lucrărilor prezentate mai sus au fost realizate folosind tehnici tradiționale pentru realizarea hărților lingvistice, respectiv acestea erau „desenate” de cartografi specializați, cel puțin în ceea ce privește transcrierile fonetice. Pentru harta propriu-zisă, ce conține granițe, localități importante, cursuri de apă și eventual formele de relief erau folosite șabloane anterior pregătite. Odată cu introducerea pe scară largă a tehnologiilor informatice în cele mai multe domenii de activitate, a apărut ideea și în anumite cazuri chiar necesitatea de a le utiliza și pentru cartografia lingvistică.

3. SOLUȚII INFORMATICE PENTRU PUBLICAREA ATLASELOR LINGVISTICE

Colaborarea dintre cercetătorii dialectologi și informaticieni a început după anul 2000 la inițiativa celor dintâi, care au început să simtă lipsa de cartografi specializați în realizarea transcrierilor fonetice pentru hărțile lingvistice și modificarea tehnologiilor utilizate în edituri și tipografii. Aceste colaborări au fost inițiate de colectivele de cercetare care se ocupă de cercetări dialectologice în Institutele Academiei Române din București, Cluj-Napoca și Iași. Noua abordare a condus, pe de o parte, la dezvoltarea de soluții informatice pentru editarea asistată de calculator a atlaselor lingvistice (problema de rezolvat pe termen scurt), dar pe de altă parte, soluțiile adoptate fiind diferite, metodele și structurile de date utilizate nu sunt unitare și necesită crearea unor aplicații de conversie, unificare și adaptare a acestora în vederea unor procesări ulterioare în comun a datelor colectate (problemă de rezolvat pe termen lung).

3.1. Sistemul de editare asistată de calculator a planșelor și textelor dialectale pentru *NALR-Moldova și Bucovina*

Primul sistem informatic pentru editarea și publicarea unui atlas lingvistic din România este (după știința noastră) cel utilizat pentru publicarea volumelor III și IV al *NALR. Moldova și Bucovina* (NALR. MB). Acesta este rezultatul colaborării dintre cercetători de la Institutul de Informatică Teoretică și Institutul de Filologie Română „A. Philippide” din cadrul filialei din Iași a Academiei Române. Colaborarea a început în anul 2000 la inițiativa lui Ion A. Florea și a avut ca obiectiv inițial crearea unei aplicații software care să permită editarea transcrierilor fonetice pentru *NALR. MB* vol. III. În cadrul programelor de cercetare ale Academiei Române desfășurate în cadrul celor două institute în perioada 2000–2005 a fost dezvoltată aplicația *ALR* (Figura 5) care a fost utilizată pentru publicarea celor două volume din *NALR. MB* în anii 2007 și 2014. Pentru stocarea informațiilor au fost proiectate și implementate două baze de date: lingvistică și geografică. Cea dintâi include 3 dicționare pentru descrierea cuvintelor titlu, punctelor de anchetă și transcrierilor fonetice, în timp ce a doua bază de date conține șablonul cu descrierea vectorială a trăsăturilor geografice care apar pe hartă dar și pozițiile implicite în care vor fi plasate titlul hărții, transcrierile fonetice, notele, etc.

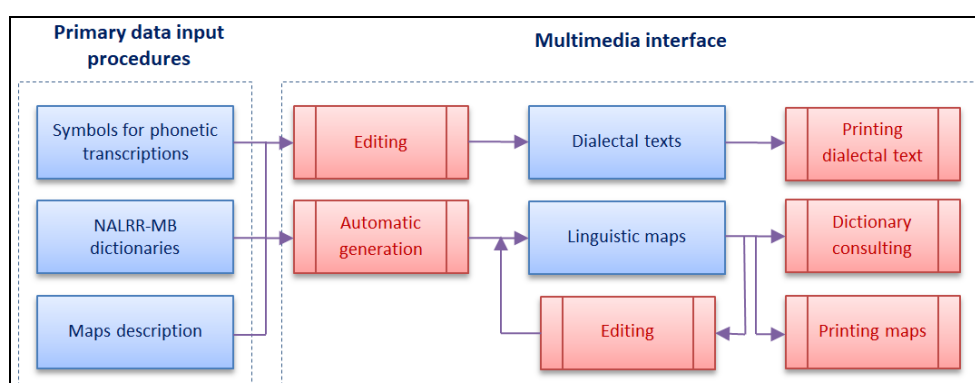


Fig. 5. Structura generală a sistemului dezvoltat pentru publicarea *NALR. MB*
(Sursa: Bejinariu, Apopei, Dumistrăcel, Teodorescu 2009: 565)

În ceea ce privește transcrierile fonetice, aplicația *ALR* permite aplicarea tuturor combinațiilor posibile de fenomene fonetice folosind doar două fonturi. În cursul editării, este necesar ca utilizatorul să selecteze simbolul de bază din tastatură (pentru unele simboluri – combinații de taste) urmată de aplicarea pe rând a fenomenelor fonetice prin selectarea acestora din bara de instrumente. Simbolurile din transcrierea fonetică sunt codificate convenabil folosind o structură de date ce include codul simbolului de bază, fenomenele fonetice aplicate, informații de poziționare și de formatare. Afișarea transcrierilor este realizată cu ajutorul unui sintetizator care generează în timp real imaginea fiecărui simbol la dimensiunea și cu proprietățile dorite (Apopei, Bejinariu, Roman 2002: 8). Dicționarul de transcrieri fonetice permite și stocarea de înregistrări audio a răspunsului primit în cursul anchetei.

Materialul dialectal stocat în baza de date poate fi valorificat prin generarea automată a planșelor cu hărți lingvistice și/sau material (ne)cartografiat. Pe aceste planșe informațiile dialectale sunt plasate în poziții prestabilite specificate în șablon și pot fi editate ulterior prin re poziționare, modificarea unor setări de afișare, dar și adăugarea de elemente grafice noi (de exemplu izoglose, imagini etc.). Sistemul permite și generarea automată a hărților sintetice și interpretative (Figura 6) cu arii delimitate, simboluri și hașuri dacă modul de grupare al punctelor de anchetă a fost stabilit în dicționarul de transcrieri (Olariu F.-T., Olariu V., Bejinariu, Apopei 2007).

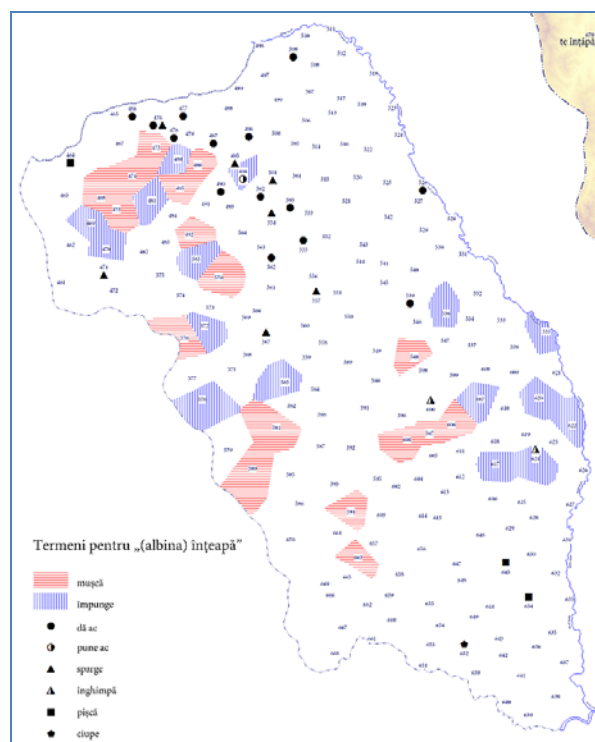


Fig. 6. Hartă sintetică din *NALR – Moldova și Bucovina*, vol. III (Sursa: imagine generată de autori folosind aplicația *ALR*)

Aplicația *ALR* include și un modul pentru editarea de texte dialectale. Folosind o interfață similară cu cea pentru editarea dicționarului de transcrieri și comenzile uzuale pentru formatare ce sunt disponibile în cele mai multe editoare de text, utilizatorul poate edita cu ușurință textele dialectale în vederea publicării lor în volum.

Deși funcționalitatea de bază a aplicației *ALR* a fost implementată în urmă cu aproximativ 15 ani, aceasta a fost îmbogățită permanent cu noi facilități și este utilizată pentru pregătirea volumului V din *NALR. MB*. Amintim între acestea posibilitatea de a genera automat pagini Web cu listele (format tabelar) cu răspunsurile din anchete. Experiența acumulată a condus la implementarea unei noi versiuni a aplicației, bazată pe conceptul de Sistem Informatic Geografic (Bejinariu, Luca, Olariu 2016: 38(2)), versiune ce este descrisă în secțiunea următoare.

3.2. Aplicația *AlrMaps* pentru publicarea *Atlasului lingvistic al dialectului aromân*

Aplicația utilizată pentru *NALR. MB* prezintă dezavantajul că descrierea hărții este codificată în aplicație ceea ce conduce la necesitatea „adaptării” programului în cazul în care se dorește utilizarea lui pentru o altă zonă dialectală. Chiar dacă această operațiune este ușor de realizat, inconvenientul este că trebuie utilizată câte o versiune a aplicației pentru fiecare hartă. Odată cu demararea colaborării dintre Institutul de Lingvistică al Academiei Române „Iorgu Iordan – Al. Rosetti” și Institutul de Informatică Teoretică în cadrul proiectului fundamental al Academiei Române pentru realizarea *Atlasului lingvistic al dialectului aromân*, în perioada 2017–2018 a fost proiectată și implementată o nouă versiune a acestei aplicații, numită *AlrMaps*. A fost utilizată experiența acumulată până în acel moment, fiind eliminate toate neajunsurile identificate în utilizarea aplicației *ALR*. Structura dicționarelor a fost îmbunătățită prin adăugarea unor noi structuri de date care permit individualizarea completă a șabloanelor planșelor lingvistice, așa cum este detaliat în (Bejinariu, Olariu 2017). În acest scop, noua aplicație a fost reproiectată folosind conceptul de Sistem de Informații Geografice (SIG). În fapt, șablonul este pregătit folosind QGIS – o aplicație gratuită open-source de tip SIG. Pot fi generate automat planșe cu harta lingvistică (cu sau fără hartă sintetică), harta cu rețeaua localităților, harta cu repartitia graiurilor precum și planșe cu material (ne) cartografiat. Harta este organizată pe straturi și prin urmare la generarea hărților pot fi alese elementele care vor fi vizibile sau nu. De asemenea, proprietățile sunt stocate la nivel de hartă (implicite) dar și pentru fiecare element în parte permițând un nivel sporit de personalizare a hărților (Figura 7). În *AlrMaps* a fost implementată și posibilitatea de a afișa o imagine de fundal ce conține relieful ariei analizate.

O componentă importantă aflată într-o fază finală de implementare este noul editor de texte dialectale pentru atlasele sociolingvistice *AlrWord AddIn*. Acesta este implementat ca *add-in* pentru editorul de text Microsoft Word (Figura 8). În bara de comenzi este adăugat un panou ce conține comenzile specifice editării transcrierilor fonetice. Principala inovație adusă este aceea că în bară sunt afișate permanent variantele simbolului de bază pe care se află cursorul, iar utilizatorul îl poate modifica oricând în timpul editării și prin urmare nu mai este necesară utilizarea de shortcut-uri sau a ferestrei de inserare simboluri pentru selectarea unui caracter. În ceea ce privește aplicarea fenomenelor fonetice, este disponibilă o interfață asemănătoare cu cea din aplicația *AlrMaps*. Utilizatorul poate selecta astfel orice combinație de fenomene fonetice. Toate simbolurile sunt organizate în 7 fonturi TrueType (în funcție de numărul de fenomene aplicate), fiind utilizată și o regulă pentru stabilirea corespondenței dintre codul caracterului și fenomene aplicate – utilă pentru alte prelucrări automate ale textului dialectal.

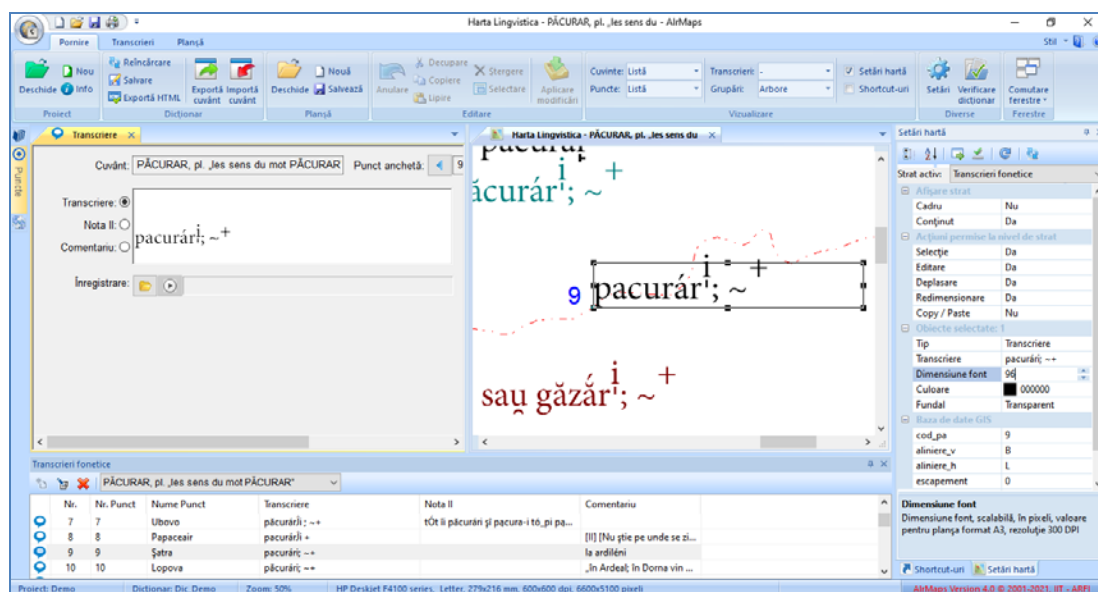


Fig. 7. Modificarea setărilor utilizate la desenarea transcrierilor fonetice (Sursa: captură din aplicația *AlrMaps* dezvoltată de autori)

Aplicația *AlrMaps* a fost utilizată pentru generarea planșelor lingvistice și pregătirea pentru publicare a *ALAR* vol. 2, de Nicolae Saramandu și Manuela Nevaci (editor).

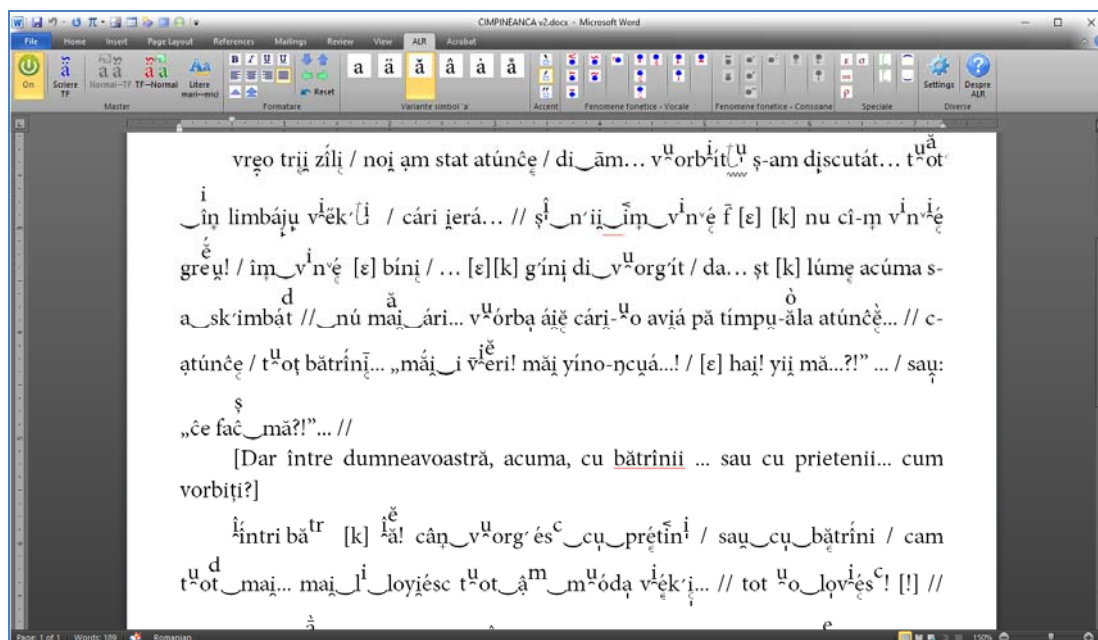


Fig. 8. Editarea transcrierilor textelor dialectale în Microsoft Word (Sursa: captură din aplicația *AlrLibrary* dezvoltată de autori, textul dialectal „651. Cîmpineanca (Florești)” din *NALR. MB. Texte dialectale*).

3.3. Aplicația RODA pentru publicarea *NALR-Crișana*

O altă abordare bazată pe utilizarea tehnologiilor informatice pentru realizarea atlaselor lingvistice este cea utilizată la publicarea volumelor III și IV ale *NALR. Crișana* în anii 2011 și respectiv 2017, fiind datorată colaborării dintre grupul de cercetare condus de Dorin Uritescu de la York University, Toronto și Institutul de Lingvistică și Istorie Literară „Sextil Pușcariu” din Cluj-Napoca. Astfel, a fost dezvoltată aplicația Romanian Online Dialect Atlas (în două versiuni, *RODA I* și *RODA II*) care, în afară de introducerea materialului dialectal în baza de date, permite: vizualizarea informațiilor, căutări în colecția de date, contorizarea aparițiilor

unui anumit șablon și reprezentarea vizuală a acestora, crearea de hărți fonetice și interpretative ce pot fi publicate în volum sau on-line, aplicarea unor tehnici de analiză (Embleton, Uritescu, Wheeler 2013: 16).

Una dintre noutățile aduse de sistemul *RODA* în geolingvistica românească este realizarea de hărți interpretative de tip dialectometric. Folosind tehnica scalării multidimensionale (MDS – multi-dimensional scaling), distanțele dintre puncte sunt calculate statistic realizând o grupare a acestora în funcție de fenomenele analizate. Un exemplu de astfel de hartă este prezentată în Figura 9. O realizare importantă este și digitalizarea în *RODA* a materialului dialectal inclus în primele două volume ale *NALR. Crișana* care au fost publicate folosind tehnologia clasică, punând la dispoziție acest material pentru atlasul online.

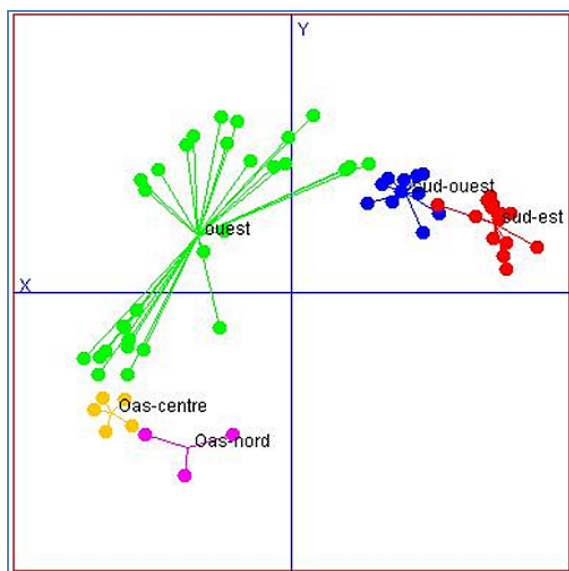


Fig. 9. Grafic MDS care separă zone distincte:
Oaș, sud-est, sud-vest, și zone intermediare
(Sursa: detaliu din Fig. 2 – Embleton, Uritescu,
Wheeler 2013: 16)

3.4. Atlasul lingvistic audiovizual al Bucovinei

O realizare deosebită în cartografia lingvistică, publicată doar în varianta web, este reprezentată de *Atlasul lingvistic audiovizual al Bucovinei* (ALAB). În cadrul a două proiecte de cercetare succesive (*ALAB* și *ALAB faza II*) implementate la Institutul de Filologie Română „A. Philippide” al Academiei Române Filiala Iași, sub coordonarea lui Florin-Teodor Olariu a fost realizat în premieră un atlas lingvistic ce utilizează sincretismul sunet-imagie video în redarea materialului dialectal (*ALAB II*). Prin intermediul interfeței web utilizatorii au acces la materialul lingvistic structurat în: răspunsuri punctuale la întrebările din chestionar și un corpus de etno- și socio-texte înregistrate de la fiecare subiect anchetat. Proiectul ALAB, care a valorificat și experiența acumulată implementarea sistemului de editare asistată de calculator a planșelor și textelor dialectale pentru *NALR. Moldova și Bucovina*, aduce și alte noutăți metodologice: (a) introducerea variațiilor diagenetice și diasexuale printre parametrii

urmăriți în anchete și (b) constituirea unei arhive digitale care să includă documente redactate în registru dialectal (Olariu F.-T., Olariu V., Clim, Luca 2016: 204). Intenția autorilor este de a transforma proiectul *ALAB* într-un „atlas-‘muzeu lingvistic virtual’, susceptibil de a facilita, astfel, abordări inovative în domeniul analizei variațiilor geo- și sociolingvistice” (*ALAB II*).

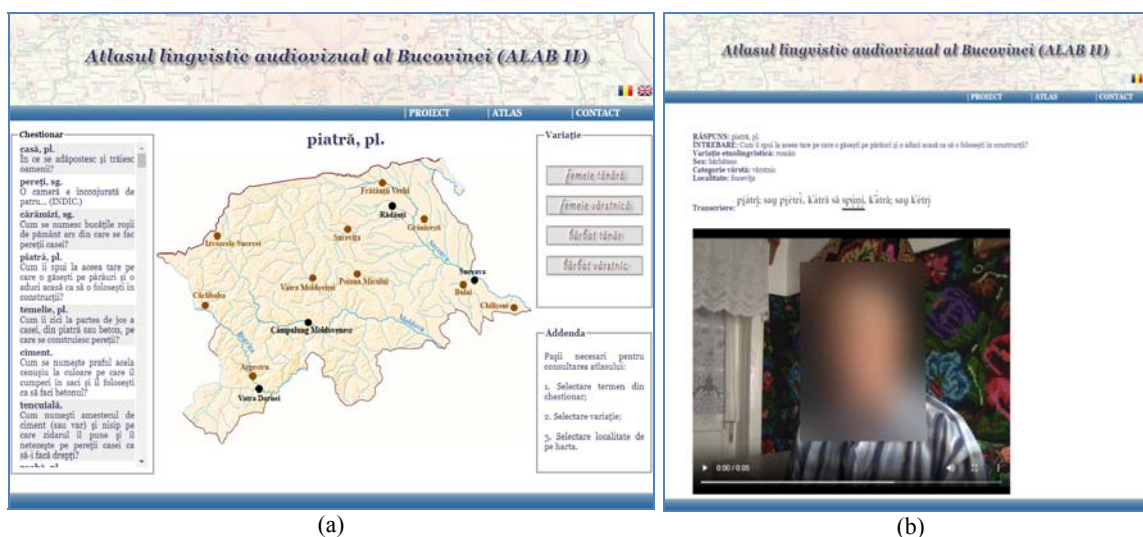


Fig. 10. Pagini web din site-ul ALAB. (a) Harta, chestionarul și meniurile de selecție a variației, (b) Pagina pentru redarea audio-video a răspunsului prezentat și în transcriere fonetică. Pe site imaginea vorbitorului nu este blurată. (Sursa: ALAB II, site-ul Atlasul lingvistic audiovizual al Bucovinei ALAB philippide.ro)

Spre deosebire de celelalte atlase lingvistice, în *ALAB* nu este prezentată o hartă lingvistică în adevăratul sens al cuvântului. Pe site este afișată doar imaginea hărții cu localitățile în care s-au efectuat anchete colorate în funcție de variația (vârsta și sexul) vorbitorului selectat (Figura 10.a), iar transcrierile asociate sunt afișate simultan cu secvența video în care este redat răspunsul vorbitorului (Figura 10.b) după selectarea punctului de anchetă pe hartă.

3.5. Alte soluții informatice pentru publicarea atlaselor lingvistice

După cum a fost deja menționat, utilizarea sistemelor informatice pentru publicarea atlaselor lingvistice a fost determinată în principal de modificarea tehnologiilor utilizate în edituri sau tipografii. O asemenea situație este întâlnită la ultimul volum al *ALRR. Muntenia și Dobrogea*. În acest caz transcrierile fonetice au fost redactate pe foi de calc în urmă cu mai mulți ani urmând ca în tipografie acestea să fie suprapuse peste șablonul hărții. Datorită modificărilor tehnologice această modalitate de tipărire nu a mai fost posibilă. În plus, foile de calc s-au degradat în timp fiind necesară o restaurare calitativă. În cadrul colaborării dintre Institutul de Lingvistică al Academiei Române „Iorgu Iordan – Al. Rosetti” (ILIR) și Institutul de Informatică Teoretică (IIT) a fost elaborată și implementată o metodă pentru recuperarea materialului și publicarea planșelor lingvistice (Bejinariu, Saramandu, Nevaci, Apopei, Teodorescu 2020: 223). Prin metode specifice prelucrării de imagini calitatea imaginilor obținute prin scanarea celor 228 de foi de calc a fost îmbunătățită folosind aplicația ImagePro (Figura 11) dezvoltată de IIT. De asemenea, a fost reproiectat șablonul hărților care a fost

suprapus în final cu imaginile prelucrate ale foilor de calc generând imaginile finale pregătite pentru tipar.

Această soluție a fost una hibridă, în sensul că planșele finale sunt disponibile în format electronic dar materialul dialectal nu este stocat într-o bază de date care să permită prelucrarea automată ulterioară ci doar în fișiere text.

Tot o soluție hibridă a fost adoptată și în cazul volumului III al *ALRR. Sinteză* coordonat de Nicolae Saramandu. Deoarece hărțile sunt redactate folosind simboluri, transcrierile fonetice ale formelor apar doar în legende, ceea ce permite editarea acestora folosind un editor de text. Codificarea formelor și asocierea cu simbolurile este realizată cu ajutorul unei aplicații de calcul tabelar (Microsoft Excel). Planșele finale sunt realizate cu ajutorul unei aplicații informatice care suprapune pe șablonul hărții legenda și simbolurile corespunzătoare. Și în acest caz materialul dialectal este cules doar în format text, la fel ca în *ALRR – Muntenia și Dobrogea*.

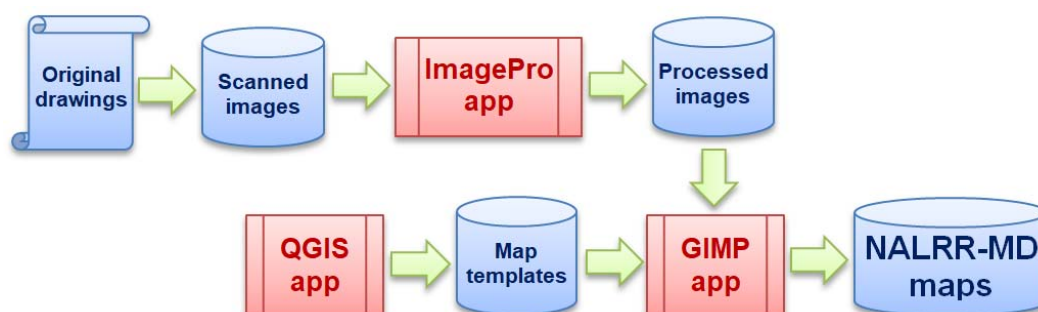


Figura 11. Structura procedurii pentru restaurare și pregătirea publicării planșelor *NALRR-MD* (Sursa: Bejinariu, Saramandu, Nevaci, Apopei, Teodorescu 2020: 232)

Pentru materialul dialectal care este disponibil în format electronic (text) există totuși posibilitatea implementării unor aplicații de conversie care să permită codificarea și stocarea acestuia într-o bază de date relațională oferind astfel nu doar posibilitatea prelucrării ulterioare dar și valorificarea acestuia prin diferite forme de prezentare. O astfel de abordare este deja implementată (Bejinariu 2019: 179(33)), tot ca rod al colaborării dintre Institutul de Lingvistică al Academiei Române „Iorgu Iordan – Al. Rosetti” și Institutul de Informatică Teoretică, pentru elaborarea hărților din *Atlas Linguarum Europae* vol. 1 (fasc. 9) – Nicolae Saramandu, Manuela Nevaci și Ionuț Geană.

În ultimii ani a fost inițiată publicarea pe internet a conținutului atlaselor lingvistice în format scanat, ceea ce oferă posibilitatea consultării lor în scop de cercetare și învățământ. Având în vedere că edițiile vechi sunt în număr redus, greu de găsit fiind disponibile doar în centrele de cercetare specializate și biblioteci, și nu în ultimul rând dimensiunile și greutatea lor, considerăm publicarea lor pe internet ca fiind o inițiativă ce vine în ajutorul celor interesați de studiul evoluției Limbii Române. Menționăm contribuția cercetătorilor de la Institutul de Lingvistică al Academiei Române „Iorgu Iordan – Al. Rosetti” care au publicat pe site-ul propriu (http://lingv.ro/index.php?option=com_content&view=article&id=381&Itemid=270): *ALAR* vol. 1 și *ALRR. Sinteză* vol. III, ambele coordonate de Nicolae Saramandu. De asemenea, pe site-ul Institutului de Lingvistică și Istorie Literară „Sextil Pușcariu” (<http://inst-puscariu.ro/al.html>) este disponibilă o colecție impresionantă ce include toate volumele din: *ALR I*, *ALR II*, *ALRM I*, *ALRM II*, *ALR II* serie nouă, *ALRM II* serie nouă, *ALRR. Maramureș*, *NALR. Banat*, *ALRR. Transilvania* și *NALR. Crișana*.

4. CONCLUZII

În ultimii 20 de ani au fost aduse contribuții importante în ceea ce privește crearea de instrumente de cercetare informatizate pentru dialectologii români. În această lucrare au fost trecute în revistă principalele rezultate obținute de colectivele institutelor de cercetare ale Academiei Române, cu accent pe sistemele software de realizare a atlaselor lingvistice și sociolingvistice realizate în cadrul colaborărilor cu Institutul de Informatică Teoretică Iași.

BIBLIOGRAFIE

- ALAB II – *Atlasul lingvistic audiovizual al Bucovinei (ALAB II)*, <http://www.philippide.ro/alab/rezumat.html>, accesat la data de 8.04.2021.
- ALAR – *Atlasul lingvistic al dialectului aromân*, Nicolae Saramandu, Manuela Nevaci (editor), vol. 1, București, Editura Academiei Române, 2014.
- Apopei, Bejinariu, Roman 2002 – Vasile Apopei, Silviu-Ioan Bejinariu, Mariana Roman, “Graphic Symbols Generator for the Phonetic Transcription in the Electronic Linguistic Atlas”, în *Proc. of ECIT'2002, European Conferences on Intelligent Systems and Technologies*, Iași, România, July 2002.
- Bejinariu, Saramandu, Nevaci, Apopei, Teodorescu 2020 – Silviu-Ioan Bejinariu, Nicolae Saramandu, Manuela Nevaci, Vasile Apopei, Horia-Nicolai Teodorescu, “Recovery of old dialectal materials and maps through image processing”, *Romanian Journal of Information Science and Technology (ROMJIST)*, 23(3), pp. 223–237, 2020.
- Bejinariu 2019 – Silviu-Ioan Bejinariu, “Information management framework for the European Linguistic Atlas”, în *Fonetică și dialectologie*, XXXVIII, pp. 147–196, 2019.
- Bejinariu, Olariu 2017 – Silviu-Ioan Bejinariu, Florin-Teodor Olariu, “Romanian Linguistic Atlases in Digital Format A New Approach”, în *Philologica Jassyensia*, 13(1), Ed. Tracus Arte, Bucharest, pp. 13–23, 2017.
- Bejinariu, Luca, Olariu 2016 – Silviu-Ioan Bejinariu, Ramona Luca, Florin-Teodor Olariu, “A GIS Based Approach for Information Management in Geolinguistics”, în *Memoirs of the Scientific Sections of the Romanian Academy*, Tome XXXIX, Computer Science, pp. 37–45, 2016.
- Bejinariu, Apopei, Dumistrăcel, Teodorescu 2009 – Silviu-Ioan Bejinariu, Vasile Apopei, Stelian Dumistrăcel, Horia-Nicolai Teodorescu, “Overview of the Integrated system for dialectal text editing and Romanian Linguistic Atlas publishing – 2009”, în *Proc. of the 13-th International Conference “Inventica 2009”*, Iași, Ed. Performantica, pp. 564–572, 2009.
- Cugno 2013 – Federica Cugno, „Legături între *Atlante Linguistico Italiano* și *Atlasul Lingvistic Român*”, în *Caietele Sxtil Pușcariu*, I, *Actele Conferinței Internaționale „Zilele Sxtil Pușcariu”*, Ediția I, Cluj-Napoca, 12–13 septembrie 2013, pp. 53–60, 2013.
- Embleton, Uritescu, Wheeler 2013 – Sheila Embleton, Dorin Uritescu, Eric S. Wheeler, “Defining dialect regions with interpretations: Advancing the multidimensional scaling approach”, în *Literary and Linguistic Computing*, vol. 28(1), pp. 13–22, 2013.
- NALR. MB. Texte dialectale – *Noul Atlas lingvistic român, pe regiuni. Moldova și Bucovina. Texte dialectale*, vol. I, partea a 2-a, culese de Stelian Dumistrăcel și publicate de Doina Hreapcă și Ion-Horia Bîrleanu, Iași, Editura Academiei Române, 1995.
- Olariu F.-T., Olariu V., Clim, Luca 2016 – Florin-Teodor Olariu, Veronica Olariu, Marius-Radu Clim, Ramona Luca, «La cartographie linguistique roumaine face à l’informatisation : quelques projets et résultats», în J.-P. Chauveau, M. Barbato, I. Fernández-Ordóñez (eds.), *Actes du XXVIIe Congrès international de linguistique et de philologie romanes (Nancy, 15–20 juillet 2013). Section 8 : Linguistique variationnelle, dialectologie et sociolinguistique*, Nancy, ATILF, pp. 193–207, 2016.
- Olariu F.-T., Olariu V. 2010 – Florin-Teodor Olariu, Veronica Olariu, „O sută de ani de cartografie lingvistică românească – un bilanț deschis.”, în *Philologica Jassyensia*, An VI, Nr. 1 (11), pp. 89–118, 2010.
- Olariu F.-T., Olariu V., Bejinariu, Apopei 2007 – Florin-Teodor Olariu, Veronica Olariu, Silviu Bejinariu, Vasile Apopei, “Los atlas lingüísticos rumanos: entre manuscrito y formato electrónico”, în *Revista española de lingüística*, Año n° 37, Fasc. 1, pp. 215–246, 2007.
- Pavel 2012 – Vasile Pavel, „Termeni pentru fierăria meșteșugărească și tâmplărie în graiurile moldovenești din zona de nord (pe baza ALRR.Bas.)”, în *Buletin de lingvistică*, (13), pp. 26–36, 2012.

- Pavel 1993 – Vasile Pavel, „Atlasul lingvistic român, pe regiuni. Basarabia, Bucovina de nord, Transnistria. Reconsiderări, sarcini, modalități de cartografiere”, al *VI-lea Simpozion Național de Dialectologie*, Cluj-Napoca, pp. 159–167, 12–14 octombrie 1990.
- Petrovici 1954 – E. Petrovici, „Repartiția graiurilor dacoromâne pe baza *Atlasului Lingvistic Român*”, în *Limba română*, III (5), pp. 5–17, 1954.
- Saramandu 2007 – Nicolae Saramandu, «Atlas Linguistiques Roumains», în *Dialectologia et Geolinguistica*, 15, pp. 90–101, 2007.
- Saramandu 1990 – Nicolae Saramandu, „Pentru un atlas al atlaselor lingvistice regionale”, în *Limba română*, XXXIX (1), pp. 57–67, 1990.
- WLAD – *Linguistischer Atlas des dacorumänischen Sprachgebietes* / herausgegeben auf kosten der rumänischen Academie in Bukarest von Prof. Dr. Gustav Weigand – Leipzig: Johann Ambrosius Barth, 1909, disponibil cu licență *Public Domain* pe site-ul Europeana, https://www.europeana.eu/en/item/9200456/handle_123456789_54079, accesat la data de 07.04.2021.

¹ *Institutul de Lingvistică al Academiei Române
„Iorgu Iordan – Alexandru Rosetti”,
Calea 13 Septembrie nr. 13, București,*

² *Institutul de Informatică Teoretică,
Academia Română, Filiala Iași,
Str. T. Codrescu, nr. 2, 700481*