

PENTRU UN ATLAS MULTIMEDIA PROZODIC ROMÂN (AMPROM)

de

Adrian TURCULEȚ, Anca BIBIRI, Oana PANAIȚE

Universitatea „Alexandru Ioan Cuza” Iași

Introducere. Inițiat încă din 2001 la Grenoble de Michele Contini, din 2004 beneficiind de concursul echipei coordonate de Adrian Turculeț pentru domeniul limbii române, proiectul AMPER este aproape necunoscut cercetătorilor limbii române. Articolul de față își propune prezentarea, în linii mari, a proiectului AMPER, din care a derivat AMPROM, stadiul actual al cercetărilor, al rezultatelor și perspectivelor pentru domeniul limbii române.

AMPER. Atlas de ultimă generație care îmbină resursele lingvistice cu instrumentele informatice, AMPER (*Atlas Multimédia Prosodique de l'Espace Roman*) a fost conceput ca o completare, în domeniul prozodiei, a *Atlasul Lingvistic Romanic (ALiR)*, cu scopul de a stabili principalele modele intonaționale romanice și a variației lor diatopice. Inițiat de Michele Contini, în 2001, în cadrul Centrului de Dialectologie al Universității Stendhal din Grenoble, în urma studiilor de prozodie derulate de acesta în mai mult de 30 de ani, a tezelor de licență, masterat și doctorat consacrate prozodiei coordonate de acesta, proiectul AMPER se desfășoară în prezent prin comitete științifice pentru fiecare domeniu romanic în parte:

AMPER

Coordonatori :

Michel Contini & Antonio Romano

AMPER-AST

Coordonator :
Carmen Muniz

AMPER-CAN

Coordonator:
Josefa Dorta

AMPER-CAT

Coordonator:
Ana Maria Fernandez Planas

AMPER-ESP

Coordonator:
Eugenio Martinez Celdran

AMPER-FRA

Coordonator:
Jean-Pierre Lai

AMPER-GAL

Coordonator:
Elisa Fernandez

AMPER-ITA

Coordonator:
Antonio Romano

AMPER-POR

Coordonator:
Lurdes de Castro Moutinho

AMPER-ROM

Coordonator:
Adrian Turculeț

Spre deosebire de celelalte atlase lingvistice, AMPER nu rămâne la Romania europeană, extinzând aria cercetărilor consacrate limbilor romanice prin includerea țărilor din America Latină.

Metodologie. Atlas lingvistic din a treia generație [6, p.206], AMPER este un atlas prozodic multimedia, electronic, pluridimensional; atlas prozodic, deoarece surprinde prozodia varietăților

PENTRU UN ATLAS MULTIMEDIA PROZODIC ROMÂN (AMPROM)

romanice, prin analiza parametrilor fizici: frecvența fundamentală, intensitatea și durata, la care se adaugă ritmul și pauza; atlas multimedia, deoarece presupune înregistrarea, prelucrarea și reprezentarea video-audio a datelor, astfel pot fi vizualizate/ ascultate fișierele de sunet și ton, curba frecvenței fundamentale (F0), curba stilizată și graficele realizate prin prelucrarea datelor acustice. AMPER este un atlas informatizat care se prezintă, pentru specialiști, sub forma unei baze de date ce poate fi accesată pe Internet fără limitări temporale sau spațiale, iar pentru nespecialiști, printr-o reprezentare cartografică a punctelor „vorbitoare”; Doar unele aspecte din multitudinea de date vor fi redată cu ajutorul „hărților vorbitoare”: se va putea asculta enunțul în rostirea reală a informatorului, însoțită de imaginea transcrierii și a conturului normalizat al prozodiei enunțului. Pe hărțile de tip clasic se vor trasa ariile modelelor prozodice rezultate în urma analizelor comparative. AMPER este și un atlas pluridimensional care permite abordări complexe, din diferite perspective, a realității lingvistice, în care datele pot relaționa direct și nelimitat pentru a surprinde variația diastratică, diafazică și diacronică a graiurilor romanice.

AMPER își propune realizarea prin anchete prozodice a trei corpusuri, necesare cercetării prozodiei la nivel lexical și post-lexical în diferite contexte stilistico-pragmatice. Corpusul fix (echivalentul chestionarului clasic) este alcătuit dintr-o listă de enunțuri specifice fiecărei limbi, structurate după aceleași criterii fonetice, sintactice și semantico-pragmatice. Corpusul fix cuprinde 45 de enunțuri realizate în modalitate declarativă și interogativă. Structura sintactică a enunțurilor este SVO, în care S și O – prezentând tipuri accentuale diferite (oxiton, paroxiton și proparoxiton) – se pot substitui și dezvolta prin determinanți adjectivali sau prepoziționali. Fiecare enunț va fi repetat de minimum trei ori, cu intonația neutră specifică varietății respective.

Corpusurile *semi-spontan* (obținut prin procedee de tip Map Task ce presupun un dialog între cei doi informatori pentru a descrie modul în care se deplasează pe o hartă ad-hoc care conține ca puncte de reper locuri și lucruri denumite (și) prin cuvintele utilizate în corpusul fix) și *spontan* (conversații libere și/sau dirijate între subiecți, mici narațiuni sau descrieri). Ultimele două corpusuri permit verificarea și completarea datelor din corpusul fix.

Enunțurile sunt înregistrate în format digital (fișiere cu extensia .wav - *Waveform Audio File Format*) și analizate acustic cu ajutorul unor instrumente software.

Procesul de analiză parcurge succesiv mai multe etape:

- Schimbarea frecvenței de eșantionare a undei sonore din 48 kHz la 16 kHz (PRAAT/ GoldWave)
- Delimitarea enunțurilor și etichetarea lor conform chestionarului utilizat;

PENTRU UN ATLAS MULTIMEDIA PROZODIC ROMÂN (AMPROM)

- Segmentarea și etichetarea vocalelor (vezi Figura 1):

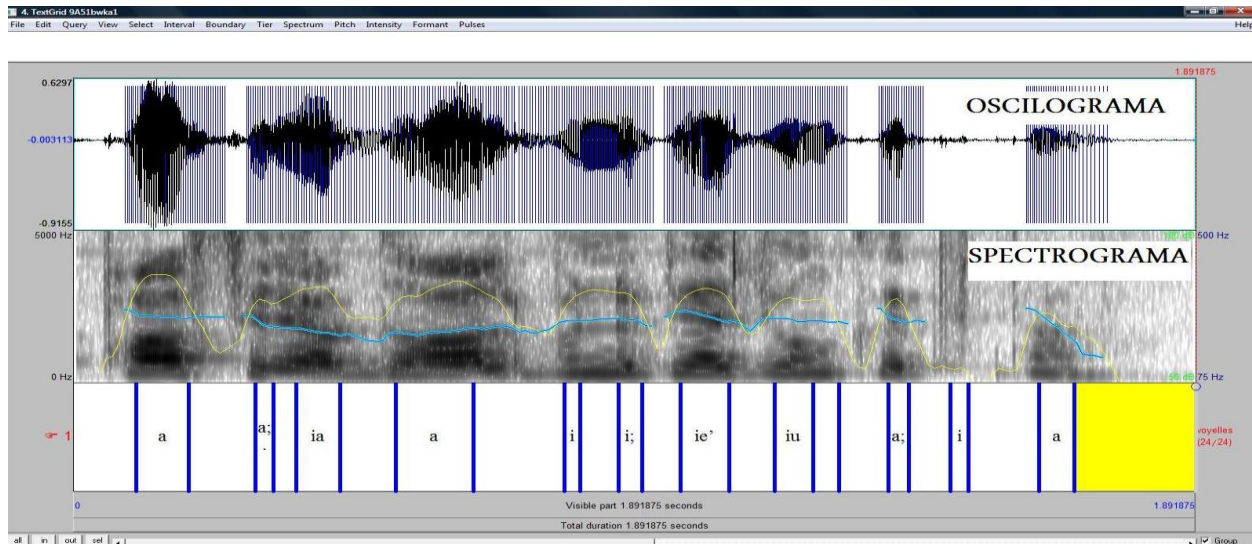


Figura 1

- - Conversia fișierului TextGrid într-un fișier .txt care va conține parametrii acustici ai fiecărei vocale: durata (ms), energia (dB) și frecvența (extrasă în câte trei puncte) – vezi Figura 2.

```
G:\AMPER\MS1\9A51bwka1.txt size: 30270
24-Apr-2010

duration [ms] energy [dB] fo1 fo2 fo3 [Hz]
1 88 85 263 259 254
2 31 77 245 236 231
3 74 82 226 216 211
4 130 81 214 217 228
5 26 78 247 248 248
6 39 80 252 250 252
7 82 81 272 262 247
8 64 78 253 249 247
9 34 78 253 250 247
10 29 55 253 254 255
11 60 71 264 234 201

values at:
1691 2394 3097 4907 5151 5396 6004 6596 7188 8691 9733 10775 13243 13449 13654 14696 15012
15327 16369 17027 17685 18919 19431 19943 21990 22264 22539 23665 23900 24135 26049 26528
27008
```

Figura 2

Analiza acustică realizată de AMPER este o analiză „maximalistă“ a parametrilor fizici prin care sunt reprezentate fenomenele prozodice: frecvența F0, intensitatea și durata vocalelor din enunț. Pe baza schemelor reprezentând evoluția în timp a F0 și a intensității (energiei sonore) obținute din analiza repetițiilor fiecărui enunț se realizează modelarea contururilor prozodice, în funcție de frecvența laringiană medie a vorbitorului și prin eliminarea micilor variații ale tonului datorate unor fenomene de microprozodie. Prin tratamentul statistic al mai multor repetiții se obțin modele de contururi globale normalizate pentru diferite tipuri de enunțuri, care păstrează trăsături fonostilistice și atitudinale și

PENTRU UN ATLAS MULTIMEDIA PROZODIC ROMÂN (AMPROM)

mărcile prozodice ale provenienței geografice a individului. Aceste contururi redată prin diferite grafice facilitează cercetarea comparativă. Acest demers are ca finalitate stabilirea modelelor prozodice specifice varietățile romanice pentru tipurile de enunțuri investigate (asertive și interogative totale) și a distribuției lor în spațiu.

AMPRom. Concomitent cu proiectul internațional, echipa de specialiști de la Iași coordonată de Adrian Turculeț a propus realizarea unui *Atlas multimedia prozodic român (AMPRom)* prin care se va cerceta tipologia prozodiei limbii române atât la nivelul variantei literare, cât și la nivelul variantelor dialectale. Atlasele mai vechi (ALR – *Atlasul Lingvistic Român*, NALR – *Noul Atlas Lingvistic Român*) nu au consemnat intonația, astfel încât realizarea unui atlas prozodic, care să cuprindă principalele modele intonaționale românești, este o necesitate. Instrumentele de lucru sunt programele de analiză realizate în cadrul AMPER (rutine de procesare a semnalelor dezvoltate de Antonio Romano și Albert Rilliard cu ajutorul Matlab) și PRAAT.

În cadrul anchetelor dialectale realizate pe teren, chestionarul AMPRom se aplică la cel puțin doi informatori (un bărbat și o femeie) în fiecare punct de anchetă. Punctele de anchetă sunt selectate, în principal, din rețeaua ALR și NALR. Informatorii trebuie să îndeplinească criteriile clasice necesare într-o anchetă dialectală: disponibilitate comunicațională, vârstă medie, voce bună, să fie originari din localitatea cercetată.

Corpusul fix. Seturile de enunțuri care compun chestionarul – stabilite după criteriile sintactice și fonetice – sunt formate din: propoziții *declarative* (afirmative și negative), propoziții *interogative* totale (afirmative și negative), cu structura sintactică SVO (subiect – verb - obiect), în care S și O primesc, pe rând, determinante adjectivale și/sau prepoziționale; substantivele și adjectivele utilizate în enunțuri sunt trisilabice oxitone, paroxitone și proparoxitone. Deoarece în limba română, negația primește, de obicei, accentul frazei, au fost introduse în chestionar și enunțuri negative:

Chestionar AMPER (secvență):

twk Nevasta vede un căpitan	kwt Un căpitan vede nevasta
dwk Nevasta tinerea vede un căpitan	gwt Un căpitan elegant vede nevasta
swk Nevasta frumoasă vede un căpitan	pwt Pasărea vede nevasta
zwk Nevasta harnică vede un căpitan	bwt Pasărea papagal vede nevasta
twg Nevasta vede un căpitan elegant	fwt Pasărea frumoasă vede nevasta

PENTRU UN ATLAS MULTIMEDIA PROZODIC ROMÂN (AMPROM)

Pentru a surprinde un număr mai mare de tipare intonaționale românești în răspândirea lor teritorială, a fost conceput un al doilea chestionar care cuprinde și alte tipuri de enunțuri, cum ar fi cele cu subiect inclus sau subînțeles de tipul VO, enunțuri enumerative și exclamative:

Chestionar AMPRom (secvență) :

(1) 1a : L-ai văzut pe Ion? 1b Pe Ion l-ai văzut?(2) 2a: L-ai văzut pe Vasile?, 2b Pe Vasile l-ai văzut? (3) 3a: Ai văzut fetele? 3b: Fetele le-ai văzut? (4) 4a: L-ai văzut pe V a s i l e ? 4b: Pe V a s i l e l-ai văzut ? (5) L-ai văzut pe Ion la biserică?

Culegerea datelor privind intonația prezintă și multe dificultăți, datorită variabilității care depinde de factori obiectivi și subiectivi, cum ar fi: tiparele intonative funcționale care caracterizează vorbitorul nativ, contextul în care se desfășoară situația de comunicare, structura sintactică a enunțului, structura ritmică a cuvintelor folosite, la care se adaugă factorii de expresivitate și cei afectivi greu de controlat.

Enunțurile trebuie obținute într-o atmosferă de familiaritate între anchetator și informatori, realizarea unui dialog pentru contextualizarea enunțurilor, sugerarea unor cuvinte cu ajutorul unor desene, schițe sau al procedeelelor de tipul Map Task (pentru corpusul s e m i l i b e r) sau chiar sugerarea lor directă în alte contexte. Informatorul nu trebuie să citească enunțurile pentru a evita intonația specifică „lecturii” Enunțurile trebuie rostite neutral, spontan, cât mai apropiat de vorbirea obișnuită a informatorului, fără focalizarea unui anumit cuvânt. Fiecare s trebuie repetat de cel puțin trei ori, în reprize diferite, cu pauze între reluări.

Corpusul l i b e r constă în înregistrarea unor conversații libere între informatori, între anchetatori și informator, povestiri ale informatorului, pentru a permite culegerea unor enunțuri cât mai variate, realizate în mod spontan care vor fi comparate cu cele din corpusul fix. Ancheta va începe cu acest tip de corpus, pentru ca anchetatorul să observe modelele intonaționale specifice și să aleagă informatori reprezentativi.

Rețeaua de puncte AMPRom. În proiectul AMPER sunt prevăzute 10 puncte de anchetă (având codurile 90-99) pentru România și Republica Moldova, majoritatea sunt localități rurale, dar și orașe, pentru varietățile diatopice ale limbii literare.

Rețeaua de puncte a AMPRom-ului pornește de la provinciile istorice și acoperă întreg teritoriul dialectului dacoromân: Moldova, Muntenia, Transilvania, Maramureș, Crișana, Banat, Oltenia, Dobrogea, Basarabia, dar și teritoriile dialectelor aromân și meglenoromân. Densitatea punctelor de anchetă va depinde de reprezentativitatea acestora. Apreciem că ar putea fi suficiente câteva zeci (cca 50-60) de puncte. Puținele cercetări acustice-auditive asupra graiurilor dacoromâne au arătat că ariile

PENTRU UN ATLAS MULTIMEDIA PROZODIC ROMÂN (AMPROM)

intonaționale nu coincid cu ariile dialectale care se pot trasa cu ajutorul datelor segmentale, prezente în atlasele lingvistice românești. AMPRom își propune cercetarea graiurilor dacoromâne din perspectiva prozodică, pentru a arăta suprapunerea și deosebirile între ariile prozodice și cele dialectale.

Realizări și perspective. În perioada de pregătire a anchetelor pentru AMPRom, la Seminarul de Fonetica și Dialectologie, cu implicarea studenților și masteranzilor Facultății de Litere, s-au realizat anchete de probă în unele localități din Moldova, Bucovina, Muntenia, Transilvania, Maramureș și Republica Moldova. S-a constituit astfel o primă bază de date pentru cercetarea variației intonației la nivelul limbii literare și a variantelor ei regionale. Datele și prelucrările lor s-au valorificat în diferite comunicări și lucrări științifice.

În prezent ne aflăm în faza de prelucrare acustică a datelor culese din anchetele dialectale realizate în Moldova și Bucovina, etapă la care participă o echipă de doctoranzi și cercetători din cadrul *Departamentului de Cercetare Interdisciplinară în Domeniul Științelor Umaniste* al Universității „Alexandru Ioan Cuza” Iași. Prin obținerea a două burse post-doctorale, sunt în desfășurare proiectele *AMPRom. Maramureș* și *AMPRom. Crișana*; unele dintre punctele anchetate fac parte din proiectul internațional AMPER.

Realizarea site-ului AMPRom va permite consultarea bazei de date pentru limba română atât de către specialiștii în domeniu, cât și de toți cei interesați de probleme de prozodie.

Concluzii. Prin volumul considerabil de date, prin caracterul deschis, prin caracteristicile specifice (*multimedia, digital, pluridimensional*), *Atlasul multimedia prozodic român* va fi un instrument de lucru indispensabil pentru cercetarea variației diatopice, dar și a unor aspecte diastratice și diafazice la nivelul prozodiei limbilor romanice.

Referințe bibliografice:

1. *Projet AMPER: Atlas Multimédia Prosodique de l'Éspace Roman*. În „Géolinguistique”, hors série 3 (ed. J.P. Lai), 2005.
2. Michel Contini. *2e Séminaire international du projet AMPER*. În *Projet AMPER*, 2005, p. I- XI.
3. Michel Contini. *Le projet AMPER: passé, present, et avenir*. În Moutinho, Coimbra, 2007, p. 9- 19.
4. Adrian Turculeț, Ana-Maria Minuț. *De la AMPER la AMPRom*. În Luminița Hoarță Cărăușu (coord.) *Rezultate și perspective actuale ale lingvisticii românești și străine*, Editura Universității „Al. I. Cuza”, Iași, 2007, pp. 349-361.
5. Adrian Turculeț (ed.), *La variation diatopique de l'intonation dans le domain roumain et roman*, Iași, Editura Universității „Al. I. Cuza”, 2008.
6. Adrian Turculeț. *Un nou atlas lingvistic romanic: AMPER*. În *Studii și cercetări lingvistice*, 2007, ianuarie-iunie, București, pp. 203-214.