



**Universitatea Națională de Știință și Tehnologie**  
**POLITEHNICA București**  
**Facultatea de TEOLOGIE, LITERE, ISTORIE ȘI ARTE**



**FIȘA DISCIPLINEI**

**1. Date despre program**

1.1 Instituția de învățământ superior	<b>Universitatea Națională de Știință și Tehnologie POLITEHNICA din București</b>
1.2 Facultatea	Teologie, Litere, Istorie și Arte
1.3 Departamentul	Departamentul de Limbi Străine Aplicate
1.4 Domeniul de studii universitare	Filologie Inginerie electronică, telecomunicații și tehnologii informaționale
1.5 Programul de studii universitare	<i>Plurilingvism și inteligență artificială</i>
1.6 Ciclu de studii universitare	Master
1.7 Limba de predare	Română
1.8 Locația geografică de desfășurare a studiilor	Pitești

**2. Date despre disciplină**

2.1 Denumirea disciplinei/	<i>Lingvistică computațională</i>						
2.2 Titularul/ii activităților de curs	Conf.univ.dr. Ilinca Cristina						
2.3 Titularul/ii activităților de seminar / laborator/proiect							
2.4 Anul de studiu	2	2.5 Semestrul	I	2.6. Tipul de evaluare	E	2.7 Statutul disciplinei	Ob <sup>1</sup>
2.8 Categoria formativă <sup>2</sup>	S	2.9 Codul disciplinei	P.M.23.F.11.II.Ob.17.				

**3. Timpul total (ore pe semestru al activităților didactice)**

3.1 Număr de ore pe săptămână	2	Din care: 3.2 curs	1	3.3 seminar/laborator/proiect	1
3.4 Total ore din planul de învățământ	28	Din care: 3.5 curs	14	3.6 seminar/laborator/proiect	14
Distribuția fondului de timp:					ore
Studiul după manual, suport de curs, bibliografie și notițe Documentare suplimentară în bibliotecă, pe platformele electronice de specialitate/Pregătire seminarii/ laboratoare/proiecte, teme, referate, portofolii și eseuri					85
Tutorat					10
Examinări					2
Alte activități (dacă există):					
3.7 Total ore studiu individual	97				
3.8 Total ore pe semestru	125 <sup>3</sup>				
3.9 Numărul de credite	5 <sup>4</sup>				

**4. Precondiții (acolo unde este cazul)**

<sup>1</sup> Obligatorie / Opțională / Facultativă – Se va completa conform planului de învățământ.

<sup>2</sup> Fundamentală / de specializare/ complementare – Se va completa conform planului de învățământ.

<sup>3</sup> Se va calcula ținând cont că se acordă un credit pentru volumul de muncă care îi revine unui student cu frecvență la zi pentru a echivala 25 de ore de pregătire pentru dobândirea rezultatelor învățării.

<sup>4</sup> Se va completa conform planului de învățământ.



**Universitatea Națională de Știință și Tehnologie**  
**POLITEHNICA București**  
**Facultatea de TEOLOGIE, LITERE, ISTORIE ȘI ARTE**



4.1 de curriculum	Parcurgerea și/sau promovarea următoarei discipline: <i>Lingvistică generală</i>
4.2 de rezultate ale învățării	Competențe minimale dobândite în domeniul lingvisticii generale

**5. Condiții necesare pentru desfășurarea optimă a activităților didactice (acolo unde este cazul)**

5.1 de desfășurare a cursului	Cursul se va desfășura într-o sală dotată cu videoproiector și computere.
5.2 de desfășurare a seminarului/laboratorului/proiectului	-

**6. Obiectiv general**

Această disciplină se studiază în cadrul domeniului *Filologie*, programul de studii *Plurilingvism și inteligență artificială* și are următorul obiectiv **general** :

- Formarea de competențe legate de utilizarea și adaptarea uneltelor de prelucrare a limbajului natural, precum și dezvoltarea capacității de a contribui la activitatea de cercetare și dezvoltare în acest domeniu.

**Obiective specifice:**

- însușirea noțiunilor fundamentale ce țin de prelucrarea limbajului natural
- familiarizarea cu metode de statistică folosite în acest domeniu
- formarea de competențe în utilizarea instrumentelor de prelucrare a limbajului natural
- formarea de competențe în utilizarea sistemele de traducere automată

**7. Rezultatele învățării**

Competențele asociate disciplinei *Lingvistică computațională* sunt corelate cu următoarele rezultate ale învățării.

**CP3** Utilizează software pentru memorii de traducere. (3 ECTS)

**CP7** Evaluează tehnologiile de traducere. (2 ECTS)

Cunoștințe	<b>C3.</b> Studentul/absolventul descrie modul de utilizare a instrumentelor CAT. <b>C7.</b> Studentul/absolventul explică funcționarea tehnologiilor care permit dispozitivelor TIC să înțeleagă și să interacționeze cu utilizatorii prin intermediul limbajului uman, inclusiv prin traducere.
Abilități	<b>A3.</b> Studentul/absolventul facilitează traducerea lingvistică eficientă, utilizând software pentru memorii de traducere. <b>A7.</b> Studentul/absolventul utilizează tehnologii de traducere și furnizează observații cu privire la utilizarea acestora în scopuri definite.



Responsabilitate și autonomie	<p><b>RA3.</b>Studentul/absolventul demonstrează autonomie și responsabilitate în utilizarea software-ului de memorii de traducere, aplicând cunoștințele lingvistice și tehnice dobândite pentru a asigura o traducere coerentă, eficientă și conformă cerințelor proiectului.</p> <p><b>RA7.</b>Studentul/absolventul dă dovadă de autonomie și discernământ în utilizarea tehnologiilor de traducere, oferind observații relevante și argumentate privind eficiența și adecvarea acestora în raport cu obiectivele specifice ale proiectului de traducere.</p>
-------------------------------	---

## 8. Metode de predare

Procesul didactic va integra o varietate de metode de predare, adaptate specificului disciplinei. Se vor utiliza atât metode expositive (precum prelegerea și prezentarea teoretică a conceptelor), cât și metode conservative și interactive, menite să încurajeze învățarea activă și explorarea critică a conținuturilor. Activitățile vor include: analiza și interpretarea resurselor lingvistice digitale și a produselor generate automat (cum ar fi traducerile automate sau textele sintetizate); explorarea directă și asistată a aplicațiilor informatice specifice procesării limbajului natural (analiză morfosintactică, recunoaștere automată a vorbirii, traducere automată, generare de text etc.); precum și metode centrate pe acțiune, cum ar fi exercițiile aplicative, proiectele practice, studiile de caz și rezolvarea de situații complexe, individual sau în echipă.

Predarea va fi susținută prin prezentări PowerPoint, materiale video, precum și prin instrucțiuni detaliate privind utilizarea software-urilor dedicate care vor fi puse la dispoziția studenților pentru activitățile practice.

Fiecare curs va debuta cu o scurtă recapitulare a noțiunilor abordate anterior, punându-se accent pe conexiunile dintre conceptele teoretice discutate și aplicabilitatea lor practică în traducerea multimedia și în analiza comunicării vizual-auditive.

## 9. Conținuturi

CURS		
Capitolul	Conținutul	Nr. ore
I	<b>Introducere în lingvistică computațională.</b> Prezentarea generală a domeniului. Noțiuni elementare de prelucrare a limbajului natural.	2
II	<b>Noțiuni elementare de statistică.</b> Aplicarea modelelor statistice în prelucrarea limbajului natural. Noțiuni introductive de statistică	2
III	<b>Modele de limbă.</b> Calculul probabilității pentru unigrame, bigrame și trigrame. Netezire statistică pentru evenimente necunoscute	2
IV	<b>Etichetare morfosintactică.</b> Prezentare generală. Metode clasice folosite în etichetare morfosintactică: Modele Markov de ordinul întâi.	2
V	<b>Analiză sintactică (parsare).</b> Parsare la nivel de text, frază. Dependență sintactică și interpretare semantică.	2
VI	<b>Unelte pentru prelucrarea limbajului natural folosite în scop academic .</b> Prezentarea uneltelor de bază în prelucrarea limbajului natural. Stanford NLP Tools. Unelte RACAI pentru limba română	2
VII	<b>Modele lingvistice mari pentru generare de text.</b> Mod de utilizare Gemini, BERT, XLNet, ChatGPT <b>Extragerea automată a echivalenților de traducere.</b> Noțiunea de corpus paralel și corpus comparabil. Metode de aliniere automată	2
<b>Total:</b>		<b>14</b>
<b>Bibliografie:</b> Boroș, T., Dumitrescu, S. D., Ion, R., Ștefănescu, D., and Tufiș, D. (2013). <i>Romanian-English statistical translation at RACAI</i> . ISSN 1843-911x, 81.		



Boroș, T., Radu, I., & Tufiș, D. 2013. Large tagset labeling with Feed Forward Neural Networks. Case study on Romanian Language. In *Proceedings of ACL*.

Hausser, R. (2014). *Foundations of Computational Linguistics*. Springer, Berlin, Heidelberg.

Ilinca, C. 2019. *Lingvistică computațională* (note de curs), suport electronic

Ion, Radu and Barbu Mititelu, Verginica and Păiș, Vasile and Irimia, Elena and Badea, Valentin (2024). A Cross-model Study on Learning Romanian Parts of Speech with Transformer Models. *Proceedings of the Sixth International Conference on Computational Linguistics in Bulgaria (CLIB 2024)*. pp. 6--13, 2024

Paul, Michael J. 2012. Mixed membership Markov models for unsupervised conversation modeling. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, EMNLP-CoNLL '12, pages 94–104, Stroudsburg, PA.

Tran, Quan Hung, Ingrid Zukerman, and Gholamreza Haffari. 2016. Inter-document contextual language model. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 762–766.

Joty, Shafiq and Tasnim Mohiuddin. 2018. Modeling Speech Acts in Asynchronous Conversations: A Neural-CRF Approach, *Computational Linguistics*, Volume 44, Issue 4, December 2018 p.859-894

Păiș, Vasile and Ion, Radu and Avram, Andrei Marius and Mitrofan, Maria and Tufiș, Dan. (2024). RELATE: a modern processing platform for Romanian language. *Tools, Methods, and Solutions for the Exploration of Romanian Corpora*. Institutul European, pp. 25—52.

Păiș, Vasile and Mitrofan, Maria and Avram, Andrei-Marius. (2022). Romanian Language Translation in the RELATE Platform. *Proceedings of the Fifth Workshop on Technologies for Machine Translation of Low-Resource Languages (LoResMT 2022)*. Association for Computational Linguistics, pp. 64—74.

Reiter, Ehud. 2018. A Structured Review of the Validity of BLEU, *Computational Linguistics*, Volume 44, Issue 3, September 2018, p.393-401

Tufiș, Dan (2022). Romanian Language Technology – a view from an academic perspective. *International Journal of Computers Communications & Control*. vol. 17.

Ungureanu, Dan and Badeanu, Madalina and Marica, Gabriela-Catalina and Dascalu, Mihai and Tufis, Dan Ioan (2021). Establishing a Baseline of Romanian Speech-to-Text Models. *2021 International Conference on Speech Technology and Human-Computer Dialogue (SpeD)*. pp. 132—138.

## Seminar

Capitolul	Conținutul	Nr. ore
I	<b>Introducere în lingvistică computațională.</b> Prezentarea generală a domeniului. Noțiuni elementare de prelucrare a limbajului natural.	2
II	<b>Noțiuni elementare de statistică.</b> Aplicarea modelelor statistice în prelucrarea limbajului natural. Noțiuni introductive de statistică	2
III	<b>Modele de limbă.</b> Calculul probabilității pentru unigrame, bigrame și trigrame. Netezire statistică pentru evenimente necunoscute	2
IV	<b>Etichetare morfosintactică.</b> Prezentare generală. Metode clasice folosite în etichetare morfosintactică: Modele Markov de ordinul întâi.	2
V	<b>Analiză sintactică (parsare).</b> Parsare la nivel de text, frază. Dependență sintactică și interpretare semantică.	2
VI	<b>Unelte pentru prelucrarea limbajului natural folosite în scop academic .</b> Prezentarea uneltelor de bază în prelucrarea limbajului natural. Stanford NLP Tools. Unelte RACAI pentru limba română	2
VII	<b>Modele lingvistice mari pentru generare de text.</b> Mod de utilizare Gemini, BERT, XLNet, ChatGPT <b>Extragerea automată a echivalenților de traducere.</b> Noțiunea de corpus paralel și corpus comparabil. Metode de aliniere automată	2
<b>Total:</b>		<b>14</b>

## Bibliografie:

Boroș, T., Dumitrescu, S. D., Ion, R., Ștefănescu, D., and Tufiș, D. (2013). *Romanian-English statistical translation at RACAI*. ISSN 1843-911x, 81.

Boroș, T., Radu, I., & Tufiș, D. 2013. Large tagset labeling with Feed Forward Neural Networks. Case study on Romanian Language. In *Proceedings of ACL*.

Hausser, R. (2014). *Foundations of Computational Linguistics*. Springer, Berlin, Heidelberg.

Ilinca, C. 2019. *Lingvistică computațională* (note de curs), suport electronic



Ion, Radu and Barbu Mititelu, Verginica and Păiș, Vasile and Irimia, Elena and Badea, Valentin (2024). A Cross-model Study on Learning Romanian Parts of Speech with Transformer Models. *Proceedings of the Sixth International Conference on Computational Linguistics in Bulgaria (CLIB 2024)*. pp. 6--13, 2024

Paul, Michael J. 2012. Mixed membership Markov models for unsupervised conversation modeling. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, EMNLP-CoNLL '12, pages 94–104, Stroudsburg, PA.

Tran, Quan Hung, Ingrid Zukerman, and Gholamreza Haffari. 2016. Inter-document contextual language model. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 762–766.

Joty, Shafiq and Tasnim Mohiuddin. 2018. Modeling Speech Acts in Asynchronous Conversations: A Neural-CRF Approach, *Computational Linguistics*, Volume 44, Issue 4, December 2018 p.859-894

Păiș, Vasile and Ion, Radu and Avram, Andrei Marius and Mitrofan, Maria and Tufiș, Dan. (2024). RELATE: a modern processing platform for Romanian language. *Tools, Methods, and Solutions for the Exploration of Romanian Corpora*. Institutul European, pp. 25—52.

Păiș, Vasile and Mitrofan, Maria and Avram, Andrei-Marius. (2022). Romanian Language Translation in the RELATE Platform. *Proceedings of the Fifth Workshop on Technologies for Machine Translation of Low-Resource Languages (LoResMT 2022)*. Association for Computational Linguistics, pp. 64—74.

Reiter, Ehud. 2018. A Structured Review of the Validity of BLEU, *Computational Linguistics*, Volume 44, Issue 3, September 2018, p.393-401

Tufiș, Dan (2022). Romanian Language Technology – a view from an academic perspective. *International Journal of Computers Communications & Control*. vol. 17.

Ungureanu, Dan and Badeanu, Madalina and Marica, Gabriela-Catalina and Dascalu, Mihai and Tufis, Dan Ioan (2021). Establishing a Baseline of Romanian Speech-to-Text Models. *2021 International Conference on Speech Technology and Human-Computer Dialogue (SpeD)*. pp. 132—138.

## 10. Evaluare

Tip activitate	10.1 Criterii de evaluare	10.2 Metode de evaluare	10.3 Pondere din nota finală
10.4 Curs	- capacitatea de asimiliare și aplicare a cunoștințelor învățate; - gradul de asimilare a limbajului de specialitate și capacitatea de comunicare pe temele cursului; - capacitatea de a opera cu cunoștințele asimilate în activități intelectuale complexe.	Evaluare formativă	
		- Referat	30%
		- Evaluare sumativă (la încheierea perioadei de studiu al disciplinei de învățământ) – probă orală	50%
10.5 Seminar/laborator/proiect	- capacitatea de asimiliare și aplicare a cunoștințelor învățate;		20%
10.6 Condiții de promovare			
Exemplu:			
<ul style="list-style-type: none"><li>• Obținerea a 50% din punctajul total.</li><li>• Obținerea a 50% din punctajul aferent activității pe parcursul semestrului:</li><li>• Cunoașterea minimală a etichetare morfosintactică și analiză sintactică (parsare).</li></ul>			



**Universitatea Națională de Știință și Tehnologie**  
**POLITEHNICA București**  
**Facultatea de TEOLOGIE, LITERE, ISTORIE ȘI ARTE**



Data completării	Titular de curs	Titular(ii) de aplicații
23.09.2025	Conf.univ.dr. Cristina Ilinca	Conf.univ.dr. Cristina Ilinca
Data avizării în departament	Director de departament	
24.09.2025	Conf.univ.dr.Laura Cîțu	
Data aprobării în Consiliul Facultății	Decan	
29.09.2025	Conf.univ.dr. Constantin Augustus Bărbulescu	