



## D3.18. Diseminare

Acstea rezultate au fost obținute prin finanțare în cadrul Programului PN-III Proiecte complexe realizate în consorții CDI, derulat cu sprijinul MEN – UEFISCDI,  
Cod: PN-III-P1-1.2-PCCDI-2017-0818, Contract Nr. 73 PCCDI/2018:  
**“SINTERO: Tehnologii de realizare a interfețelor om-mașină pentru sinteza text-vorbire cu expresivitate”**

© 2018-2020 – SINTERO

Acest document este proprietatea organizațiilor participante în proiect și nu poate fi reprobus, distribuit sau diseminar către terți, fără acordul prealabil al autorilor.

Denumirea organizației participante în proiect	Acronim organizație	Tip organizație	Rolul organizației în proiect (Coordonator/partener)
<b>Institutul de Cercetări Pentru Inteligență Artificială “Mihai Drăgănescu”</b>	ICIA	UNI	CO
<b>Universitatea Tehnică din Cluj-Napoca</b>	UTCN	UNI	P1
<b>Universitatea Politehnica din București</b>	UPB	UNI	P2
<b>Universitatea "Alexandru Ioan Cuza" din Iași</b>	UAIC	UNI	P3

**Date de identificare proiect**

Număr contract:	PN-III-P1-1.2-PCCDI-2017-0818, Nr. 73 PCCDI/2018
Acronim / titlu:	<b>„SINTERO: Tehnologii de realizare a interfețelor om-mașină pentru sinteza text-vorbire cu expresivitate”</b>
Titlu livrabil:	<b>D3.18. Diseminare</b>
Termen:	<b>Noiembrie 2020</b>
Editor:	<b>Mircea Giurgiu (Universitatea Tehnică din Cluj-Napoca)</b>
Adresa de eMail editor:	<b>Mircea.Giurgiu@com.utcluj.ro</b>
Autori, în ordine alfabetică:	<b>Mircea Giurgiu, Beata Lorincz, Maria Nuțu, Adriana Stan</b>
Ofițer de proiect:	<b>Cristian STROE</b>

**Rezumat:**

În acest livrabil se prezintă articolele științifice susținute și publicate de către partenerul UTCN la conferințe internaționale în anul 2020, paginile web cu demonstratoare online privind adaptarea sistemului de sinteză text-vorbire la stilul și expresivitatea unui nou vorbitor, precum și modul în care au fost implicați studenții în tematica proiectului pe durata elaborării proiectelor de diplomă sau participării la stagii de practică. Inițial au fost identificate conferințele de interes din domeniul aferent cercetărilor. Vom include lista articolelor și un rezumat cu principalele rezultate. Accesul la publicații (cf. 30.11.2020) este asigurat pe pagina web a proiectului în secțiunea dedicată. De asemenea, există o serie de pagini web cu demonstrații online de voci sintetice cu vorbitorii din noul corpus Swara 2.0, din sistemele de adaptare la noi vorbitori, din sistemele generate cu date atipice, respectiv din interfața interactivă online.

**Cuprins**

1. Introducere.....	4
2. Identificarea posibilităților de publicare pe anul 2020 și realizări.....	4
3. Publicații științifice în anul 2020.....	4
4. Lucrări de licență în legătura cu tematica proiectului .....	6
5. Stagii de practică pentru studenți .....	7
6. Pagini web ale proiectului SINTERO .....	7
7. Pagini Wiki interne grupului de cercetare cu mostre audio generate de diferite versiuni ale sistemelor de sinteză text – vorbire implementare în etapa a III-a. ....	8
8. Concluzii .....	10

## 1. Introducere

Acest livrabil prezintă o sinteză a articolelor publicate în anul 2020, pagina web a proiectului și o serie de demonstrații online pentru adaptarea vocii sintetice la stilul și expresivitatea unui nou vorbitor, respectiv adaptarea rapidă a vocii sintetice folosind date audio atipice. Aceste rezultate demonstrează că prin utilizarea arhitecturilor de rețele neuronale profunde de tip Tacotron și DC-TTS se obține o calitate înaltă a vocii sintetice pentru vorbitori în limba română, calitate comparabilă cu rezultate de dată foarte recentă raportate la conferința internațională Interspeech 2020 și la competiția internațională Voice Conversion Challenge 2020. De asemenea, se prezintă referințe către paginile web ale demonstrațorilor online și tematici de cercetare în care au fost implicați și studenți pe durata stagiului de practică sau a elaborării proiectelor de diplomă.

## 2. Identificarea posibilităților de publicare pe anul 2020 și realizări

Conform cu strategia de diseminare inclusă în formularul de aplicație, pentru fiecare an calendaristic s-au identificat posibilitățile de diseminare și de publicare de articole la conferințe științifice sau în jurnale.

Pentru anul 2020 s-au identificat următoarele posibilități de publicare la conferințe internaționale:

- 2020 ISCA Interspeech, 25-29 Octombrie, Shanghai, China (<https://interspeech2019.org/>) - online
- 2020 IEEE 16th International Conference on Intelligent Computer Communication and Processing, Cluj-Napoca, 3-5 Sept 2020, Cluj-Napoca, Romania ([www.iccp.ro](http://www.iccp.ro))
- The 15th Edition of the International Conference on Linguistic Resources and Tools for Processing Romanian Language, 14-16 Decembrie, Bucuresti, Romania (<https://profs.info.uaic.ro/~consilr/>) – online
- IEEE International Conference – Intelligent Systems, 28-30 August 2020, Varna, Bulgaria (<https://www.ieee-is.org/>) - online
- 24th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, 16-18 Septembrie 2020, Venetia, Italia (<http://kes2020.kesinternational.org/>) - online
- 11th Nordic Conference on Human-Computer Interaction, Octombrie 2020, online.

## 3. Publicații științifice în anul 2020

Autori	Beata Lorincz, Maria Nuțu, Adriana Stan, Mircea Giurgiu
Titlu	„An Evaluation of Postfiltering for Deep Learning Based Speech Synthesis with limited data” ( <a href="https://ieeexplore.ieee.org/abstract/document/9199932">https://ieeexplore.ieee.org/abstract/document/9199932</a> )
Ref.	IEEE 10th International Conference on Intelligent Systems (IS), 28-30 August 2020 (online)

Rezumat	<p>Recently, deep neural network (DNN) based speech synthesis achieved close to human speech quality and became the state-of-the art in the field of text-to-speech (TTS) synthesis systems. However, a major part of its efficiency comes from the use of large quantity of high-quality speech recordings. When this data is not available, other approaches are still preferred. This paper evaluates the DNN-based postfiltering of the synthesised speech as a means to increase the quality of DNN based TTS systems trained on very limited speech resources. 20 different systems are compared objectively using the Mel Cepstral Distortion measure. The systems differ in terms of: training data, network architecture, and training method. Out of the 20 initial systems, 7 are evaluated subjectively in listening tests performed for two different speakers. Results show that even when starting from as little as 5 minutes of speech recordings, the postfiltering process improves the quality of the synthetic speech output. So it can, therefore, be used as a training strategy for TTS systems where sufficient high-quality data is not available.</p>
---------	--

Rezumat	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 10%;">Autori</td><td>Beata Lorincz</td></tr> <tr> <td>Titlu</td><td>„Concurrent Phonetic Transcription, lexical stress assignment and syllabification with deep neural networks” (<a href="https://www.sciencedirect.com/science/article/pii/S1877050920318366">https://www.sciencedirect.com/science/article/pii/S1877050920318366</a>)</td></tr> <tr> <td>Ref.</td><td>Proceedings of the 24th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, KES 2020, 16-18 Septembrie 2020, online.</td></tr> </table>	Autori	Beata Lorincz	Titlu	„Concurrent Phonetic Transcription, lexical stress assignment and syllabification with deep neural networks” ( <a href="https://www.sciencedirect.com/science/article/pii/S1877050920318366">https://www.sciencedirect.com/science/article/pii/S1877050920318366</a> )	Ref.	Proceedings of the 24th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, KES 2020, 16-18 Septembrie 2020, online.
Autori	Beata Lorincz						
Titlu	„Concurrent Phonetic Transcription, lexical stress assignment and syllabification with deep neural networks” ( <a href="https://www.sciencedirect.com/science/article/pii/S1877050920318366">https://www.sciencedirect.com/science/article/pii/S1877050920318366</a> )						
Ref.	Proceedings of the 24th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, KES 2020, 16-18 Septembrie 2020, online.						
Rezumat	<p><i>This paper evaluates four different sequence-to-sequence deep neural network architectures aimed to jointly solve the tasks of: phonetic transcription, lexical stress assignment and syllabification. These text processing tasks are considered essential components for high quality text-to-speech or automatic speech recognition systems, with the phonetic transcription being the most frequently used in these types of applications. Although each of the tasks has been individually and extensively analyzed in the scientific literature, there are few studies which target a concurrent solution for them. In general, the lexical stress assignment and syllabification are used as augmenting input features to the phonetic transcription model and not considered as target features. The proposed network architectures include recurrent, convolution and attention neural layers and were evaluated on hand-checked English and Romanian datasets. The accuracy of the models was evaluated in terms of accuracy for the concurrent prediction of all three tasks, as well as by discarding the syllabification or lexical stress predictions. The best results were obtained with a combination of convolution and attention layers, where the accuracy of the joint prediction for the three tasks was of 58.96% for English and 86.64% for Romanian. The same model for English obtains an accuracy of 59.70% when syllables are discarded and 64% when the prediction of lexical stress is ignored. With the same best performing model for Romanian an accuracy of 88.83% without syllables and 93.84% without lexical stress is obtained.</i></p>						

Rezumat	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 10%;">Autori</td><td>Adriana Stan</td></tr> <tr> <td>Titlu</td><td>„RECOApy – Data Recording, Pre-processing and Phonetic Transcription for End-To-End Speech-Based Applications” (<a href="https://www.isca-speech.org/archive/Interspeech_2020/pdfs/1184.pdf">https://www.isca-speech.org/archive/Interspeech_2020/pdfs/1184.pdf</a>)</td></tr> <tr> <td>Ref.</td><td>ISCA International Conference Interspeech 2020, 25-29 Octombrie 2020, online.</td></tr> </table>	Autori	Adriana Stan	Titlu	„RECOApy – Data Recording, Pre-processing and Phonetic Transcription for End-To-End Speech-Based Applications” ( <a href="https://www.isca-speech.org/archive/Interspeech_2020/pdfs/1184.pdf">https://www.isca-speech.org/archive/Interspeech_2020/pdfs/1184.pdf</a> )	Ref.	ISCA International Conference Interspeech 2020, 25-29 Octombrie 2020, online.
Autori	Adriana Stan						
Titlu	„RECOApy – Data Recording, Pre-processing and Phonetic Transcription for End-To-End Speech-Based Applications” ( <a href="https://www.isca-speech.org/archive/Interspeech_2020/pdfs/1184.pdf">https://www.isca-speech.org/archive/Interspeech_2020/pdfs/1184.pdf</a> )						
Ref.	ISCA International Conference Interspeech 2020, 25-29 Octombrie 2020, online.						

Rezumat	<p><i>Deep learning enables the development of efficient end-to-end speech processing applications while bypassing the need for expert linguistic and signal processing features. Yet, recent studies show that good quality speech resources and phonetic transcription of the training data can enhance the results of these applications. In this paper, the RECOApy tool is introduced. RECOApy streamlines the steps of data recording and pre-processing required in end-to-end speech-based applications. The tool implements an easy-to-use interface for prompted speech recording, spectrogram and waveform analysis, utterance-level normalisation and silence trimming, as well grapheme-to-phoneme conversion of the prompts in eight languages: Czech, English, French, German, Italian, Polish, Romanian and Spanish. The grapheme-to-phoneme (G2P) converters are deep neural network (DNN) based architectures trained on lexicons extracted from the Wiktionary online collaborative resource. With the different degree of orthographic transparency, as well as the varying amount of phonetic entries across the languages, the DNN's hyperparameters are optimised with an evolution strategy. The phoneme and word error rates of the resulting G2P converters are presented and discussed. The tool, the processed phonetic lexicons and trained G2P models are made freely available.</i></p>
---------	--

Autori	Kristen M Scott, Simone Ashby, Adriana Stan
Titlu	"Designing a Synthesized Content Feed System for Community Radio" ( <a href="https://dl.acm.org/doi/abs/10.1145/3419249.3420177">https://dl.acm.org/doi/abs/10.1145/3419249.3420177</a> )
Ref.	Proceedings of the 11th Nordic Conference on Human-Computer Interaction, October 2020
Rezumat	<i>The use of text-to-speech to generate radio content is largely unexplored, despite the importance of radio in remote parts of the world, where TTS offers a robust means of transforming data into media for low-literate audiences and those without regular internet access. How suitable are TTS voices for meeting the expectations of radio listeners and what type of content are these voices best suited to deliver? We present an application for generating automated daily synthesized weather forecasts for selected locations and language varieties, based on the provision of a regularly updated weather data service. We present results from a pilot listener study aimed at exploring people's reactions to this and other synthesized audio content, as we begin to explore best practices around the design of a synthesized content feed system for community radio.</i>

Articole acceptate spre publicare - Dan Oneață, Alexandru Caranica, Adriana Stan, Horia Cucu, "An Evaluation Of Word-level Confidence Estimation For End-to-end Automatic Speech Recognition", IEEE Spoken Language Technology Workshop, 19-22 ianuarie 2021, Virtual.

Articole în jurnale ISI (în proces de finalizare)

- Beáta Lőrincz, Elena Irimia, Adriana Stan, "RoLEX: An extended Romanian lexical dataset and its evaluation for predicting concurrent lexical information", va fi trimis către IEEE Signal Processing Letters în perioada imediat următoare.
- Beáta Lőrincz, Adriana Stan, Mircea Giurgiu, "RoNNA: Romanian neural network API", va fi trimis către IEEE Signal Processing Letters în perioada imediat următoare.

#### 4. Lucrări de licență în legătura cu tematica proiectului

- Ștefana Cîmpean - "Recunoașterea emoțiilor din vorbire folosind învățarea automată", lucrare de diplomă, iulie 2020.
- Andreea Sarca - "Automatic speech recognition system for Romanian using Deep Speech", lucrare de diplomă, iulie 2020.
- Roxana Marcu - "Automatic language identification from text", lucrare de diplomă, iulie 2020.
- Florin Ciotlăuș - "Music analysis using BLSTM and CNNs", lucrare de diplomă, iulie 2020.
- Cătălin Avram - "Automatic speaker recognition from SWARA corpus", lucrare de diplomă, iulie 2020.

## 5. Stagii de practică pentru studenți

- Georgiana Săraru - „Detecția stărilor depresive pe baza analizei semnalului vocal”, iulie-august 2020.
- Vlad Crehul - „Implementation of a Tacotron-based text to speech synthesis system”, iulie-august 2020.
- Vlad Crehul - „Testing experiments with Deep Speech automatic speech recognition for Romanian”, iulie-august 2020.
- Bogdan Oros - „Linear regression applied for speech classification”, iulie-august 2020.
- Ana Gheorghiu - „Analysis of prosodic events for music classification”, iulie-august 2020.

## 6. Pagini web ale proiectului SINTERO

<https://speech.utcluj.ro/sintero/>

The screenshot shows the main navigation bar with links for EN, DESCRIERE, CONSORTIU, ECHIPA, REZULTATE, and CONTACT. Below the navigation is a large banner with the text "SINTERO: Tehnologii de realizare a interfețelor om-mașină pentru sinteza text-vorbire cu expresivitate". The banner contains several lines of text in Romanian, likely describing the project's goals and funding. At the bottom of the banner, there are three yellow buttons labeled COBILIRO, TEPROLIN, and TADARAV.

The screenshot shows the 'REZULTATE' (Results) section of the website. It features a large yellow header with the text "Rapoarte științifice". Below this, there are three sections listing scientific reports for different project phases:

- Raport științific etapa 1 - sinteză (2018) [pdf]**
  - D1.15. Identificare pattern-uri prozodice [pdf]
  - D1.16. Metode de clasificare a stilului de exprimare din text [pdf]
  - D1.17. Analiza metodelor de control și adaptare automată a expresivității [pdf]
  - D1.18. Implementarea modulului de control automat al prozodiei [pdf]
  - D1.19. Diseminare [pdf]
- Raport științific etapa 2 - sinteză (2019) [pdf]**
  - D2.15. Implementarea modulului de identificare a stilului de vorbire și nivelului de expresivitate din analiza textului [pdf]
  - D2.16. Implementarea unui modul de adaptare la un nou vorbitor a sistemului de sinteză [pdf]
  - D2.17. Implementarea unui modul de transplantare a prozodiei unui vorbitor în sistemului de sinteză [pdf]
  - D2.18. Îmbunătățirea componentei de modelare și control al prozodiei; activități de testare, validare / demonstrare module software [pdf]
  - D2.19. Diseminare [pdf]
- Raport științific etapa 3 - sinteză (2020) [pdf]**
  - D3.15. Dezvoltarea unei noi tehnologii pentru adaptarea vocii sintetice la stilul și expresivitatea unui nou vorbitor [pdf]
  - D3.16. Dezvoltarea unei noi metode de adaptare rapidă a vocii sintetice folosind date audio atipice [pdf]
  - D3.17. Integrare tehnologie nouă și demonstrare în realizarea interfețelor om-mașină pentru sinteza text – vorbire. [pdf]
  - D3.18. Diseminare [pdf]

**Publicații**

- Beáta Lörincz, Maria Nutu, Adriana Stan, Mircea Giurgiu "An Evaluation of Postfiltering for Deep Learning Based Speech Synthesis with Limited Data", IEEE 10th International Conference on Intelligent Systems (IS), Bulgaria, 2020 [\[pdf\]](#)
- Beáta Lörincz, "Concurrent phonetic transcription, lexical stress assignment and syllabification with deep neural networks", Proceedings of the 24th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems KES2020, 2020 [\[pdf\]](#)
- Adriana Stan, "RECOApy: Data Recording, Pre-Processing and Phonetic Transcription for End-to-End Speech-Based Applications", In Proceedings of the Interspeech, Shanghai, China, 2020 [\[bib\]](#) | [\[pdf\]](#)
- Kristen M Scott, Simone Ashby, Adriana Stan "Designing a Synthesized Content Feed System for Community Radio", Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society, Estonia, 2020 [\[pdf\]](#)
- Adriana Stan, "Input Encoding for Sequence-to-Sequence Learning of Romanian Grapheme-to-Phoneme Conversion", In Proceedings of the 10th IEEE International Conference on Speech Technology and Human-Computer Dialogue (SpeD), Timisoara, Romania, 2019 [\[bib\]](#) | [\[pdf\]](#)
- Beáta Lörincz, Maria Nutu, Adriana Stan, "Romanian Part of Speech Tagging using LSTM Networks", In Proceedings of the IEEE 15th International Conference on Intelligent Computer Communication and Processing, Cluj-Napoca, Romania, 2019 [\[bib\]](#) | [\[pdf\]](#)
- Maria Nutu, Beáta Lörincz, Adriana Stan, "Deep Learning for Automatic Diacritics Restoration in Romanian", In Proceedings of the IEEE 15th International Conference on Intelligent Computer Communication and Processing, Cluj-Napoca, Romania, 2019. [\[bib\]](#) | [\[pdf\]](#)
- David A. Braude, Matthew P. Aylett, Caoimhín Ó Loide-Kemp, Simone Ashby, Kristen M. Scott, Brian O’Raghallaigh, Anna Braudo, Alex Brouwer, Adriana Stan, "All Together Now: The Living Audio Dataset", In Proceedings of Interspeech, Graz, Austria, 2019. [\[bib\]](#) | [\[pdf\]](#)
- Adriana Stan, Mircea Giurgiu, "A Comparison Between Traditional Machine Learning Approaches And Deep Neural Networks For Text Processing In Romanian", in Proc. of the 13th International Conference on Linguistic Resources and Tools for Processing Romanian Language, 22-23 November, Jassy, Romania [\[bib\]](#) | [\[pdf\]](#)

◦ D3.18- Diseminare [\[pdf\]](#)

**Prezentări video ale articolelor diseminate la conferințe cu desfășurare virtuală:**

- Beáta Lörincz, Maria Nutu, Adriana Stan, Mircea Giurgiu "An Evaluation of Postfiltering for Deep Learning Based Speech Synthesis with Limited Data", IEEE 10th International Conference on Intelligent Systems (IS), Bulgaria, 2020
- Beáta Lörincz, "Concurrent phonetic transcription, lexical stress assignment and syllabification with deep neural networks", Proceedings of the 24th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems KES2020
- Adriana Stan, "RECOApy: Data Recording, Pre-Processing and Phonetic Transcription for End-to-End Speech-Based Applications", In Proceedings of the Interspeech, Shanghai, China, 2020

## 7. Pagini Wiki interne grupului de cercetare cu mostre audio generate de diferite versiuni ale sistemelor de sinteză text – vorbire implementare în etapa a III-a.

Mostre audio aferente noilor voci din SWARA 2.0 sunt disponibile la adresa: <https://gitlab.utcluj.ro/speech/tts-samples/-/wikis/home>.

<https://gitlab.utcluj.ro/speech/tts-samples/-/wikis/Sisteme%20Tacotron2%20voci%20noi>

### Sisteme Tacotron2 voci noi

Voice ID	Sample 1 -"Tezele ar putea fi eliminate în acest an școlar, iar Ministerul Educației analizează mai multe variante."	Sample 2 - "Astfel încât să fie aplicate aceleași reguli și pentru copiii care învață în scenariul verde,"
DLL		
DOL		
EME		

<https://gitlab.utcluj.ro/speech/tts-samples/-/wikis/Date%20atipice>

- Sistem Tactron2 antrenat cu date preluate din segmentarea și transcrierea automată a unei resurse audio colectate de P1-CoBiLiRo, realizată în P3-TADARAV

Voice ID	Natural	Sample 1 - "Tezele ar putea fi eliminate în acest an școlar, iar Ministerul Educației analizează mai multe variante."	S a s
COB	▶ 0:00 / 0:07 ━━━━ ⏪ ⏴ ⏵	▶ 0:00 / 0:07 ━━━━ ⏪ ⏴ ⏵	▶

- Sistem Tactron2 antrenat cu date din audiobook-ul Mara folosind doar transcrierea ortografică (Mara-letters) sau transcrierea fonetică și accentul (Mara-phoneAcc) rezultate din rețelele de predicție concurrentă a informației lexicale

Voice ID	Natural	Sample 1 - "Tezele ar putea fi eliminate în acest an școlar, iar Ministerul Educației analizează mai multe variante."	S a s
MARA-letters	▶ 0:00 / 0:06 ━━━━ ⏪ ⏴ ⏵	▶ 0:00 / 0:08 ━━━━ ⏪ ⏴ ⏵	▶

<https://gitlab.utcluj.ro/speech/tts-samples/-/wikis/Sistem%20DC-TTS%20Date%20expresive%20neexpresive>

Model	Epoci	Sample 1	Sample 2
Forma ortografică & Neexpresiv	3000	▶ 0:00 / 0:03 ━━━━ ⏪ ⏴ ⏵	▶ 0:00 / 0:04 ━━━━ ⏪ ⏴ ⏵
Forma ortografică & Neexpresiv & Expresiv	4500	▶ 0:00 / 0:10 ━━━━ ⏪ ⏴ ⏵	▶ 0:00 / 0:05 ━━━━ ⏪ ⏴ ⏵
Forma ortografică & Neexpresiv & Expresiv	6000	▶ 0:00 / 0:10 ━━━━ ⏪ ⏴ ⏵	▶ 0:00 / 0:05 ━━━━ ⏪ ⏴ ⏵

<https://gitlab.utcluj.ro/speech/tts-samples/-/wikis/Sistem%20DC-TTS%20Date%20lexicale>

**Text sintetizat:**

Sample 1: Ce se va întâmpla cu aceasta după ce vom ieși din starea de urgență.

Sample 2: Elevii vor putea intra în școală doar cu declarația pe propria răspundere completată de către părinți.

Referință (audio natural):



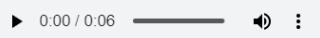
Date text	Sample 1	Sample 2
Forma ortografică	▶ 0:00 / 0:03 ━━━━ ⏪ ⏴ ⏵	▶ 0:00 / 0:06 ━━━━ ⏪ ⏴ ⏵
Forma transcrișă fonetic	▶ 0:00 / 0:04 ━━━━ ⏪ ⏴ ⏵	▶ 0:00 / 0:06 ━━━━ ⏪ ⏴ ⏵

<https://gitlab.utcluj.ro/speech/tts-samples/-/wikis/Sistem%20DC-TTS%20Englez%C4%83%20Singur%20Vorbitor%20Date%20lexicale>

## Sistem DC TTS Engleză Singur Vorbitor Date lexicale

### LJSpeech Voce

Referință (audio natural):



#### Text sintetizat:

Sample 1: The demands on the President in the execution of His responsibilities in today's world are so varied and complex

Sample 2: and the traditions of the office in a democracy such as ours are so deep-seated as to preclude absolute security.

Date text	Sample 1	Sample 2
Forma ortografică		
Forma transcrisă fonetic		
Forma transcrisă fonetic cu accent		

<https://gitlab.utcluj.ro/speech/tts-samples/-/wikis/Sistem%20DC-TTS%20Englez%C4%83%20Date%20lexicale>

## Sistem DC TTS Engleză Date lexicale

### LibriTTS voci

#### Text sintetizat:

Sample 1: Jimmie Dale slipped his mask into his pocket, and, with the parcel under his arm, stepped to the door and unlocked it.

Sample 2: He paused for an instant on the threshold for a single, quick, comprehensive glance around the room-then passed on out into the street.

Epoci	Sample 1	Sample 2
1000		
3000		

#### Voci masculine

Date text	Epoci	Sample 1	Sample 2
Forma ortografică	1000		

## 8. Concluzii

Accesul la publicațiile elaborate în anul 2020 este asigurat la adresa <http://speech.utcluj.ro/sintero/rezultate/>. Pagina web are un conținut dinamic, adaptat cu realizările din proiect, astfel că pentru această raportare se pot accesa și mostre cu semnal sintetic generat de modulul de adaptare la un noi vorbitori, așa cum este prezentat în paginile web cu demonstrațoarele.