الجمهورية الجزائرية الديمقراطية الشعبية
PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
وزارة التعليم العالي والبحث العلمي
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
جامعة ابن خلدون تيارت
IBN KHALDOUN UNIVERSITY 'TIARET'
كلية الآداب واللغات
FACULTY OF LETTERS & LANGUAGES
قسم اللغة والآدب العربي
DEPARTMENT OF ARABIC LANGUAGE & LITERATURE
مخبر الدراسات النحوية و اللغوية بين التراث و الحداثة في الجزائر
LABORATORY OF GRAMMATICAL & LINGUISTIC STUDIES BETWEEN HERITAGE & MODERNITY IN ALGERIA

PRFU research project tagged with:
Systems of Arabic language levels, from human study to automated processing
and

The laboratory-affiliated PRFU Projects Union for Linguistic Studies cooperates with the ***Faculty of Computer Science and Mathematics***
Organize

A National Conference Entitled

Epistemology of the structures of the Arabic language Between the perceptions of the human mind and the outputs of large language models

18th & 19th November 2024

"Al-Wiam Amphitheatre located in the Faculty of Letters & Languages will be hosting the event

2024-2025

Preamble: Given the fast-paced progress of information and communications technology, the Arabic language has recently regained its prominent position. This is due to its compatibility with the automated natural languages processing (**NLP)**. The phonological, morphological, syntactic, semantic, and even lexical structures of the Arabic language align well with such automated processing. This compatibility has helped Arabic regain its position, which it held for many centuries of applications and programmes used in various fields, including chatbots, sentiment analysis, machine translation, and text generation.

The Arabic language is known for its logical nature, which makes it entirely compatible with logic. Ibn Jinni and Al-Khalil have previously discussed the characteristics of the Arabic language and its correspondence with external beings. Due to this correspondence, modern grammatical and morphological tags have enabled machines to represent humans realistically. However, there are still limitations to these tags when it comes to Big Data databases. Arabic is relied upon to be one of the basic languages due to its richness in various structures that allow for detailed descriptions of situations.

The advancements in artificial intelligence and Big Data analysis techniques have paved the way for a better understanding of the Arabic language. This forum aims to shed light on the second part of the components of linguistic awareness of the Arabic tongue, building on the first part that was previously addressed in the national forum titled "Ontology of the Structures of the Arabic Language between the Philosophical Backgrounds of Human Study and Automated Processing Algorithms". The goal is to analyse the language structure in innovative and effective ways, and explore new horizons for deeper insights.

This forum discussed the physical manifestation of concepts in the Arabic language, specifically in terms of linguistic structures. The focus was on two different perspectives: human understanding and the automated processing of natural language. The second national forum is entitled "The Epistemology of the Structures of the Arabic Language: Bridging the Gap between Human Perception and Large Language Model Outputs."

In studying the structures of Arabic language, we can explore how the human mind perceives them at a linguistic level. This can be done through a comparative

analysis with large language models (LLM), which use algorithms based on formal logic to automatically process language. However, these algorithms may not yet match the level of human representation. While natural language processing (NLP) algorithms are inspired by human cognitive mechanisms, there are inherent limitations to the transfer from human to machine. Researchers are working to reproduce these cognitive mechanisms in machines, but their understanding and processing capabilities remain different from those of the human brain.

Despite the limitations, Natural Language Processing (NLP) algorithms attempt to automatically represent linguistic functions such as syntax, semantics, and language nuances that are processed by the brain. However, they are still the most advanced techniques available today and cannot fully reproduce the complexity and precision of human understanding. Researchers are encouraged to adopt an interdisciplinary approach in designing their research projects to solve various problems, which enhances our understanding of complex issues. For instance, some researchers have attempted to propose models of empathy to give a quasi-human touch to machine language representation and pave the way for more sensitive human communication. Significant progress has been made in psycholinguistic diagnosis thanks to the use of linguistic programming libraries. Neuro-Arabic tools like AraBERT and mBERT are crucial in understanding the linguistic differences and psychological aspects of the Arabic language.

A single researcher or a single specialty cannot complete this particular achievement. The topic under discussion is too broad and falls under the category of interdisciplinary sciences. It requires a combination of linguistic, psychological, social, mathematical, religious, and information skills to gain a comprehensive understanding of human language and its relevant aspects within each field of specialization.

As a child learns to speak, a variety of factors come into play, such as verbal training, psychological monitoring, social bonds, value awareness, and moral alertness. All of these components must be considered when developing an automated language treatment system. In this regard, a child's language acquisition process can be compared to the use of natural language processing algorithms. Even though the mechanisms of these processes differ, their goal is the same: to understand and produce natural language. While the similarities between man and machine may seem strange, a closer

look at a child's language acquisition process reveals that the brain undergoes a complex process that even scientists struggle to understand. However, this should not stop us from simulating these mental processes through various stages of natural language processing algorithms, such as exposure to language, learning language structures, error correction, understanding context, and expanding vocabulary and language skills.

The objectives of the forum are varied, but the most important ones are:

1. Analyzing the structures of the Arabic language: this involves identifying the mechanisms of the Arabic language structures by studying linguistic texts from different periods, and benefiting from ancient and modern linguistic theories.

2. Understanding the way humans perceive language: this involves studying the relationship between the way humans comprehend the language and its structures, and the perceptions of the human mind.

3. Utilizing large linguistic databases: this involves analysing vast amounts of linguistic data available online and in linguistic databases, to extract patterns and provide new insights about the structures of the Arabic language.

4. Implementing practical procedures of machine intelligence techniques: this involves activating artificial intelligence techniques at the structural levels of language, such as machine learning and text analysis, to develop models and tools that help understand and analyse the structures of the Arabic language.

---

[1] ***AraBERT*** *developed by Antoun et al.,(2020) is a widely adopted model pre-trained on an extensive corpus of Modern Standard Arabic (MSA) texts. AraBERT is applied in various natural language processing (NLP) tasks, including text classification, named entity recognition (NER), and sentiment analysis (SA) in the Arabic language.*

Reference*: Wissam Antoun, Fady Baly, and Hazem Hajj. 2020.* ***AraBERT****: Transformer-based model for Arabic language understanding. In Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools, with a Shared Task on Offensive Language Detection, pages 9–15, Marseille, France. European Language Resource Association.*

[1] ***mBERT*** *released by Devlin et al.,(2019) is a single-language model that was pre-trained using monolingual corpora in 104 languages, including Arabic. This enabled BERT to learn and generalize across multiple languages.*

Reference: *Jacob Devlin, Kenton Lee, Kristina Toutanova, Jacob Devlin, and Ming-Wei Chang. 2019.* ***Bert****: Pretraining of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.*

5. Developing programming algorithms that understand language contexts: this involves building smart applications and effective tools that mimic human understanding of the Arabic language, to achieve several goals such as teaching and learning the language, and automatic translation of meanings rather than just words.

6. Accelerating scientific research in academic institutions: this involves generating results and conclusions that contribute to scientific research in the fields of computational linguistics, data science, and their applications to the Arabic language, in order to keep pace with technological developments outside the walls of academic institutions.

The expected outcomes of the conference are as follows:

1. By carefully analyzing and describing the structures of the Arabic language through the latest artificial intelligence technologies and large databases, we can gain a better understanding of it and move towards automated processing of the language. This will reveal linguistic patterns and rules that were unclear before, especially ones that have been controversial in the past.

2. We aim to develop analytical and educational tools that can help researchers and Arabic language enthusiasts better understand its structures.

3. The knowledge gained from this work will improve the accuracy and efficiency of machine translation systems.

4. The study also expects to improve and develop applications related to human-machine communication, such as social robotics and linguistic constructions.

Overall, this work will open new horizons for understanding and applying Arabic language structures more deeply and effectively, leading to innovative solutions in multiple fields.

Conference Topics:

1. Analysis of the Structures of the Arabic Language:

This topic involves studying the fundamental linguistic structures in the Arabic language through the analysis of sentence structures in classical Arabic. The purpose is to identify fixed and stereotyped structures as well as explore patterns and exceptions.

2. Perceptions of the Human Mind:

This topic involves studying the linguistic processes within the human mind and how the linguistic structure is assimilated against the external physical reference. The analysis is done by examining the relationship between language and the processes of thinking, memory, and learning.

3. Big Data Analysis:

This topic involves organizing the huge linguistic data available through linguistic thesauri and using analysis techniques such as machine learning and data mining to identify engineering patterns of Arabic linguistic structures.

4. Applications of Artificial Intelligence in the Arabic Language:

This topic involves building machine-learning models to analyse linguistic structures and applying artificial intelligence techniques to develop tools for machine translation, e-learning, and text analysis.

5. Practical and Industrial Applications:

This topic involves developing practical applications that use research results to improve machine translation, developing tools for teaching the Arabic language, and other applications to facilitate linguistic communication between humans and machines.

6. Social Influence and Applications:

This topic involves assessing the social and cultural impact of artificial intelligence applications in the field of the Arabic language. The focus is on exploring how technology can be used to enhance communication and understanding between different cultures through the Arabic language.

Framing the forum:

Honorary Chairman of the Forum: **Prof. Dr. BELGOUMANE Berrezoug**, Head of Ibn Khaldoun University of Tiaret

General Supervisor of the Forum: **Prof. Dr. ZERROUKI Abdelkader**, Dean of the Faculty of Letters and Languages

Forum Chairman: **Prof. Dr. BENDJELLOUL Mokhtar**, Director of the Laboratory of Grammatical and Linguistic Studies between Heritage and Modernity in Algeria

Chairman of the Automated Processing Committee: **Dr. OUARED Abdelkader**, Lecturer A, Faculty of Computer Science and Mathematics

Chairman of the Scientific Committee: **Prof. Dr. BOUHENNOUCHE Fatima**, Member of the Grammatical and Linguistic Studies Laboratory

Chairman of the Linguistic Organization and Follow-up Committee: **Prof. Dr. BELKACEM Aissa**, Head of the Syntactical and Morphological Studies Team / Studies Laboratory

Members of the scientific committee

Prof. Arabi Ahmed ..... …..... University of Tiaret

Prof. Belhocine Mohamed..... .....University of Tiaret

Prof. Bencherif Mohamed .... University of Tiaret

Prof. Boulakhras Mohamed ....University of Tiaret

Prof. Hadouara Omar .... University of Tiaret

Prof. Hamidani Aissa. ..... University of Tiaret

Prof. Baloul Ahmed ..... ……. University of Tiaret

Prof. Belkacem Benaouda..... University of Tiaret

Dr. Belkenichi Ali ....University of Tiaret

Prof. Farez Fatima..... ..... University of Tiaret

Prof. Mis Souad..... ..... University of Tiaret

Prof. Djebbali Fatiha..... ..... University of Tiaret

Dr. Haji Zoulikha .... University of Tiaret

Dr. Yahiaoui Amer ....University of Tiaret

Prof. Kacem Kada ....University of Tiaret

Prof. Kerrache Benkhaoula ..... ….University of Tiaret

Prof. Bouzian Ahmed .... University of Tiaret

Prof. Benfraiha Djillali .... University of Tissemsilt

Prof. Gharbi Bekai .... University of Tissemsilt

Prof. Boughari Fatima..... ..... University of Tissemsilt

Prof. Belmihoub Hind .... University of Tissemsilt

Prof. Hadouara Mohamed ....Aflo University Center

Dr. Hamza Boudjmel .... Aflo University Center

Dr. Djebbari Mohamed ..... . Khamis Miliana University

Prof. Boutechente El-Ousfoura..... University of Blida 02

Prof. Oueld Enbia Youcef.....… University of Mascara

Prof. Benammar Mehieddine ....University of Batna

Prof. Brahimi Ahmed .... University of Djelfa

Dr. Belkhairi Abdelmalek .... University of Djelfa

Dr. Bouzidi Mohamed .....University of Tiaret

Dr. Harath Mohamed .....University of Chlef

Organizing and Follow-up Committee:

Prof. Benabed Ammar (English language) ..... University of Tiaret

Prof. Hemaidia Mohamed (English language) ..... University of Tiaret

Prof Malki Benaid (French language) ..... University of Tiaret

Dr. Mehdi Amir (French) ..... ……..... University of Tiaret

Dr. Soudani Mohamed (German language) ..... University of Tiaret

Dr. Brahim Khaled (German language) ..... University of Tiaret

Dr. Belarbi Khaled (English language) ..... University of Tiaret

Dr. Hemaidia GHlamallah (English language) ..... University of Tiaret

Dr. Boubakeur Abed (Spanish language) ..... University of Tiaret

Dr. Medjeddad Adda (Arabic language)..... Relizane University

Dr. Ayad Amel (Arabic language) ..... . University of Tiaret

Dr. Kalbaza Youcef (Arabic language)..... University of Tiaret

Dr. Djellali Ali... (Arabic language).....Relizane University

Dr. Benssassi Belguendouz (Arabic language).....Sidi Bel Abbes University

Dr. Messabih Larbi (Arabic language)..... University of Tiaret

Dr. Ammar Hamou (Arabic language)..... ..... University of Tiaret

Ms. Oudia Noura (Spanish language) ..... University of Tiaret

Doctoral students (class of 2023) project members PRFU

Scientific subcommittees

- By specialty, based on the forum

<table>
<tr><td colspan="3">Arabic Language Committee
Prof. Kerrache Benkhaoula - Tiaret - Prof. Bouzian Ahmed - Tiaret - Prof. Kacem Kada - Tiaret -
Mr. Dr. Farez Fatima - Tiaret - Prof. Djebbali Fatiha. Tiaret - Dr. Hadji Zoulikha. Tiaret -
Dr.. Boughari Fatima - Tissemsilt - Dr. Belmihoub Hind - Tissemsilt -</td></tr>
<tr><td><u>French Language Committee</u>
Prof. Malki Benaid- Tiaret
Dr.Mehdi Amir -Tiaret
Mr. Zouatnia Samir - Khmis-Miliana -
Mr. Benkrama Chahir, Khmis-Miliana</td><td rowspan="2"><u>Computer Science Committee</u>
Prof. Chadli Abdellatif - Tiaret
Dr.. Talebi Omar - Tiaret -
Prof. Belarbi Mustapha - Tiaret -
Dr. Merati Mdjeddad - Tiaret -</td><td>English Language Committee
Prof. Benabed Ammar - Tiaret
Prof. Hemaidia Mohamed - Tiaret
Dr. Belarbi Khaled - Tiaret -
Dr. Hemaidia Ghlamallah - Tiaret -</td></tr>
<tr><td><u>Spanish Language Committee</u>
Dr. Boubakeur Abed- Tiaret
Ms. Oudia Noura, - Tiaret</td><td><u>German Language Committee</u>
Dr. Soudani Mohamed- Tiaret
Dr. Brahim Khaled - Tiaret</td></tr>
</table>

Participation Conditions:

- The topic of the article should be relevant to one of the forum's topics.
- The research intervention should comply with the requirements specified for academic scientific research.
- The research should be original and not copied from any previous work.
- The research should be written in Traditional Arabic font size 14. Footnotes and references should be in the same font and size 12.
- References should be placed at the bottom of each page, and the insertion will be done automatically.

General instructions:

- Research will be overseen by a scientific committee.
- The Scientific Committee reserves the right to decline any research without giving reasons for rejection.
- The Forum may publish contributions in its affiliated magazine or a special book.
- Once a researcher has submitted their article, they cannot participate in another scientific demonstration or modify and publish it as an article.

- The researcher will be notified of the acceptance of their article via email or phone.

NOTE:

If you wish to participate, please fill out the participation form and send it, along with the abstract and a brief CV, to the Chairman of the Scientific Committee via the forum's email (mentioned below) no later than May 28, 2024

Important Dates:

- The deadline for submitting abstracts is **May 28th, 2024**
- The date for response to the abstracts is **June 4th, 2024**
- The deadline for receiving complete research is **October 4th, 2024**
- The date for responding to accepted research and sending invitations is **October 18th, 2024**
- The forum will be held on **November 18th /19th, 2024**

If you have any queries, please send an email to:
epistemology.structures.ar.18.11.24@gmail.com.
https://chat.whatsapp.com/EMQdSaHt6SR4H5BP5SwPwx

Participation form

picture
(of your choice)

Name and Surname: . ...

Qualification: ....

Academic rank: ....

General specialization: ....

Specialty: ....

University of Affiliation: ....

Professional email: ....

Phone Number: ....

The Axe of the intervention: .....

Title of the intervention: ....

Article title in English: ....

……

Abstract: ....

....

....

....

....

....

Key words: ....

Abstract of the intervention in English:

....

....

....

....

....

....

....