The deep integration of artificial intelligence technology in the teaching and learning of the Arabic language[1]

Prof. Dr. Khaireddine Bacha

University of Mo-nastir, Tunisia

khairi.bacha@gmail.com

Abstract

This work is located at the confluence of two areas of research, NLP and CALL, which remained long separated, while their intersections have a field of interdisciplinary research very promising. We have proposed the TELA environment in the aim to offer teachers interactive tools to generate educational activities more diversified interesting.

---

[1] - This paper was presented at the 3rd International English Language Department Conference on "AI & Technology in Languages, Humanities, Social Sciences, and Education," held from March 6th to 7th, 2024, at Al-Noor University College in Mosul, Iraq.

The integration of the techniques of NLP in the CALL offers CALL systems new capabilities of treatments that allow him to manipulate knowledge explicit and calculable, to interact with the learner.

The techniques NLP provide learning tools dynamics and the Methods of Evaluation of performance, they also allow to make evolve the learning methods and programs of teaching of the Arabic language

Keywords: TELA, NLP, CALL, pedagogical activities, Arabic language

Introduction

The Recent Evolution of Technologies of Information and Communication for the Teaching, has led the Community Computer Assisted Language Learning (CALL), to be interested in a comprehensive manner to the engineering of the multifaceted learning and distance. The CALL has benefited from the computer revolution, caused by the emergence of multimedia computers. In fact, the sites of learning online multiply. They bring a gain in terms of time and flexibility of use, in addition, a significant interactivity.

With the computer progress, the field of the automatic processing of natural language (NLP) has experienced a large growth for most of the languages. For the Arabic, the tools from the domain of NLP, have been slow to integrate in the platforms of learning languages. The number of these tools, which are put at the disposal of the CALL for offer to teachers and learners of interactive environments for learning the Arabic language, rest of our point of view, very modest and complete enough [25]. In effect, the field of this study, is very broad, there are more and more research and technologies that are concerned about the specificities of this

language [16] [19], [20], [21], [22], [23] and which offer tools necessary for the development of its automatic processing. We then propose our environment of the Arabic TELA.

The proposed approach is based on the integration of tools NLP, mainly our morphological analyzer, multifunction dictionary and spell checker in the CALL systems for the realization automatic and semi- pedagogical applications. An idea which will bring a solution to the problems of learning platforms in conventional line.

Critical of existing and the problem of interoperability

Two problematics intersect in our work: The first is the outcome of the NLP, the second moves closer to the IT problems in the CALL. Let us start by, first, describing the problematic in NLP and subsequently in CALL.

Problem of pre-existing tools NLP

The achievement of the tools NLP, according to specific objectives, is a complex and costly operation to implement. It is for this reason that it becomes paramount, benefit of language resources already developed, in order to catch up with the technological gap, in terms of content and services.

In fact, when trying to establish a tool NLP for any language from zero, as is the case of Arabic, we can take advantage of the pre-existing tools. Indeed, when looking more closely, we have found that it is very difficult, if not impossible, to exploit the existing tools for several reasons:

▪ Unusable of several tools because they are in the form of a prototype (or incomplete), non-portable, non-modifiable, Non-reusable and/or programd by old languages [28]. However, there are several morphological analyzers. Several studies have shown that the most used are the analyzers Aramorph, Al Khalil and BuckWalter (Bama). They have many weak points:

first, they do not identify the pattern of the word which limits its use in a syntactic analysis. Then, they do not use the diacritic signs which are inserted generally to reduce the ambiguity and the number of possible solutions. Finally, the output of analysis is in the format translitéré and therefore is not exploitable directly by other applications [18].

- Unavailability of most of these tools because of the lack of literature which concerns, their inexistences on the market and/ or the inability to reach the filmmaker who often, is no longer interested in the field [26].

- Some tools are limited to one or two language levels and the possibility of the reuse or the extension is difficult and expensive. The limits and theim perfection that can be criticized in these tools are numerous, such as the circularity, errors and inconsistencies [24].

- Expensive because several are paid, with unsatisfactory results; in addition they are asking for a long time to be able to adapt to our needs, without being sure of achieving the goal, because most of the tools are made by well specific applications [16].

- Inadequacy of the amount of information that contain these tools.

Problem of Arab CALL

Computer Assisted Language Learning (CALL) is the result of development of computer tools and their integrations in the teaching of languages, except that these tools present several problems. They exist a few applications for the learning of the Arabic language, presenting several disadvantages which include:

- The interfaces are very poorly made, the absence of educational tasks (the learner can select any Level and any exercise without aid and have no purpose of learning), the limited number of

available courses, the absence of automatic correction and feedback (nor automatic correction, nor aid in case of need) [28].

- The rigidity of the design of the activity (activities are written by hand and may not be easily modified or enriched) and the rigidity of the treatment of responses (the correct answers are predefined and any difference, compared to the response of the learner, even with a space in too much or a focus in more or less, will be considered incorrect) [16].

- Another problem, presented with these learning tools, is the non-Adaptability of courses with learners. In effect, the succession of exercises is independent of the answers provided. Therefore, there is no adaptation of activities to the language skills of the learner. This adaptation should be based on the detection and analysis of errors, allowing the production of feedbacks adapted [19].

The software needs of learning languages Assisted by Computer are more and more large, whatever as an aid to the teaching in class or as a tool of autonomous learning. The solution will be the integration of the techniques of NLP (automatic processing of languages) in the CALL. These techniques offer the system of new capabilities of treatment which allow him to manipulate explicit knowledge calculable and on the field of learning and to interact with the Learner [21].

Need of the NLP in the CALL for the Arabic language

Antoniadis defines in [4] a tutorial of languages as being *"...before any pedagogical product, a program that implements a Teaching solution for a problem of didactics of languages without altering, neither the solution, nor, a fortiori, the problem"*. While it can be seen that most of the current systems do not have techniques ensuring the taking into account of the properties of the language.

The solution will be the integration of the techniques of NLP in the CALL. These techniques offer the system new capabilities of treatments that allow him to manipulate knowledge explicit and calculable on the field of learning, and to interact with the Learner [17].

To ensure the integration of the NLP in the CALL, one must make visible, not only to the learner, but also to the teacher, the complexity of the formalisms of the NLP, long regarded as an obstacle to this integration and binding for the users of its systems. Traditional applications of CALL suffer from some limitations due to the fact that they treat the linguistic data as a simple string of characters, while the language is a system of forms and concepts [4]. This difference of point of view between computer scientists and educationalists of languages has caused a negative impact on the traditional applications of CALL at several levels which will be detailed by the following:

- The rigidity of the software: First, during the design of the activities that are written by hand and cannot be easily modified or enriched. Then, at the level of the transformation of responses, the good answers are predefined and any difference is in the response of the learner, even, a space in too much or a focus in more or less will be considered incorrect

- The non-Adaptability of learner's courses: Two types of courses are usually offered by the software of CALL [2]: a linear route, whose activities are proposed in a predefined order and a free course where the learner chooses the order of processing activities. Only in the two cases, the succession of exercises is independent of the provided answers. Therefore, there is no adaptation of activities to the language skills of the learner. This adaptation should be based on the detection and analysis of errors allowing the production of adapted feedbacks [16].

For the domain of computer assisted language learning, Arabic language always knows a delay compared to the other European languages. We are doing the same observation: No tools to learning Arabic as the foreign language are sufficiently completed.

Contribution of the NLP to the CALL

The automatic processing of the language is more and more in development and applications of treatments increase now a days. But, for the Arabic language, it is still insufficient because of the complexity of the Arabic morphology and the irregular structures of this language. In addition, the deficiency of the processing tools of the Arabic does not prevent to create environments of aid based on the good integration and employment The NLP tools available with a depth improvement.

Surely, the integration of tools NLP to the Arabic language is more profitable than to devote themselves to improve the systems of learning the Arabic language with low performance rates and low appearances on the Internet network. Also, this smart integration facilitates the implementation, controlled correctly. First, the major constraints are to assemble these two technologies (NLP and CALL), focused on the performance of the tools NLP which does not reach acceptable percentages to be integrated in systems of learning. Those who cause immediate results of erroneous treatments, where they are unusable for the learning process and it is unacceptable to give erroneous information to the learner. Secondly, the availability of these tools NLP is very limited. They are developed in universal laboratories, funded by the Ministry of Research; which causes a difficulty of use and to obtain a tool NLP. Thirdly, the integration of tools NLP, within an application of learning requires technical skills and knowledge in software engineering as well as technicians of NLP for the implementation of these technologies [27].

In the framework of the learning of languages, the operation of each level of this system must be considered, detailed, manipulated and put in practice in a targeted manner; the links between the two levels must be highlighted specifically, and the polysemy, source of difficulties in learning of languages, must be able to find the place and the appropriate treatment. Only the use of the NLP allows, currently, to expect to achieve this goal, create systems and tools enabling teachers of languages to manipulate the language such that it is structured, in their discipline and not as the basic techniques of it are able to comprehend. As teachers of languages, a system of learning of languages may not be valid, or even acceptable, that if he is able, on the one hand to generate only the language knowledge correct, and on the other hand, to assess correctly the language productions of learners [4].

Another problem which is present: it is very difficult to integrate these existing tools on the platform MIRTO (Multi-interactive programming by research on texts and oral) [2] realized within the laboratory LIDILEM[2], where there is already a search engine to didactic orientation that allows you to find texts, following educational criteria [4]. The platform achieved is a reference to the integration of tools NLP in the learning of several languages by computer. It demonstrates the solutions that the NLP can bring to these environments [3]. It is the origin of the formulation of our work of learning the Arabic language.

For this, we must develop our own tools NLP, to be able to find solutions that appear in different levels of the automatic processing of languages. This has motivated, even more, to realize our learning environment TELA " Toward Environmental Learning Arabic "[5]. It promotes the acquisition of lexical the Arabic language. This environment, focuses on many problems related to the use of the

---

[2] Http://w3.u-grenoble3.fr/lidilem/labo/web/presentation.php

words in Arabic. First, it allows teachers to language the creation of learning activities and evaluate, to ensure sequences of cooperation, to diagnose finely some linguistic mechanisms [3], [4], [15], [17]. Then, it allows you to lead a number of automatisms and of communicative skills necessary to master of the Arabic language. Finally, there is a need to develop tools NLP that would attend the learner effectively in the resolution of the problems facing it [26].

Has the discovery of TELA: environment conducive to learning the Arabic language

Functional description

TELA is a learning environment, including solutions from the NLP. It is easy to use and allows teachers of Arabic, non-specialists of the NLP, to use resources, to set in order to generate activities aimed at Learning.

This environment will use tools from the NLP to create educational activities enabling learners to learn Arabic.

These tools of automatic processing of languages will be used in the generation of educational activities for the teaching of the Arabic [13]:

- A segmentation tool [10].
- A tool of morphological analysis" TELAMA "capable to provide the different interpretations of each form analyzed [12].
- A tool for automatic translation on the Pair Arab/Arabic [6].
- A tool for automatic translation on the Pair Arabic/English [8].
- A tool for automatic translation on the Pair Arabic/French [9].
- A tool for diversion of the arabic [7].
- A tool of conjugation of the arabic [11].
- A tool of spelling correction [14].

We are in chains with the description of the educational activities that you can generate with TELA. Finally, we conclude this chapter with a description of the features and the technical aspects of TELA.

The same resources used to create activities in TELA will also be integrated into another platform complete enough for the learning of languages. This project piloted by Georges Antoniadis is known under the name MIRTO (Multi-interactive programming by research on texts and oral). The Platform carried out Allows you to provide teachers of language the possibility of designing teaching scenarios in benefiting fully of tools NLP learners for the learning of several languages [3]. This has motivated even more to achieve TELA.

The TELA environment has been developed following a principle of modular design ensuring a maximum of flexibility and scalability. This approach was not in fact an optional choice but rather a necessity justified by two points: The first is that the system is too complex to be represented by a single module and the second is that the system is subject to possible extensions that can affect several levels, since the acquisition of the texts and up to the display of results.

Level of integration of tools NLP in TELA

The integration of the NLP in the CALL resides in the multidisciplinary work between computer scientists, note makers of languages and specialists of the NLP. In effect, one of its main objectives of our work is to provide learners of Arabic language tools adapted to their problematic didactic. This development application tools to NLP large language coverage and performance. For the realization of our environment for learning the Arabic TELA, we had recourse to our multifunction dictionary and a tool

of morphological analysis with its resources described in the previous chapters.

The integration of the morphological analyzer TELAMA provides teachers a wide choice of interactive activities. With the tools NLP that we have and the other resources, we have the opportunity to offer to teachers and learners an environment fostering a learning objective and a multitude of activities, depending on the levels. For the creation of these activities, the teacher has an ergonomic interface which obscures the treatments of generations of activities. The creation of an activity allows generating the exercises from any text, a fortiori from the texts of a corpus. Allow teachers to choose the media of these activities according to criteria under their problem becomes a necessity and this project begins to the suite by the launch of the project TELA. This learning environment of the Arabic language assisted by computer, which covers many issues related to the use of the words in Arabic, based on tools of NLP whose purpose is to provide an important educational function by allowing the learner to discover information beyond the scope of the expression of the year. Activities proposed by TELA.

The Environment TELA works based on tools NLP, allowing to create multimedia activities in a didactic perspective. These tools, commonly referred to as generators of activity, build self-correcting exercises[3] (gap exercises, remission in order, multiple-choice quizzes…). At the outset, we are as well just a few texts that we have labelled by information of order morpho-syntactic in the aim of resolving the problems of lexical access and understanding of texts and add the different meaning of the missing words in our dictionary. Then we turn to build learning activities and assess. It is to manipulate so automatic or semi-automatic language content. The whole problem of the assessment and the feedback is in the

---

[3] That is to say authorizing an automatic correction

taking into account of different levels of variability in the wrong answers. For this analysis, we propose a new approach that allows us to apply the basic techniques allowing comparison between response given (RD) / expected answer (RA), in order to obtain an analysis of differences, the output of which will then be used for the production of a feed-back, in the form of a multifunction dictionary which is composed of several NLP tools based on an architecture multi-agent [13]. The usefulness of these tools to help are to assist a learner in the process of doing its activities in providing him with morphological rules such that the Arabic translation - English, the Arabic translation - French, conjugation and the bypass… to discover the grammatical category of some words, used by the learner to improve and deepen its knowledge [3].

Feedback: Integration of multifunction dictionary in TELA

The feedback is not limited to statements of type "C is correct" or "it is false", but must above all, take other forms as links on tools NLP (conjugation, translation bypass …) within which the subject may have a summary of the conduct of the activity. These linguistic resources available helped the learner to understand texts, to the retention of the vocabulary and to the mastery of grammar. This will help us to improve the learning environment of the Arabic realize.

Fig.2. Feedback: integration of language resources in TELA
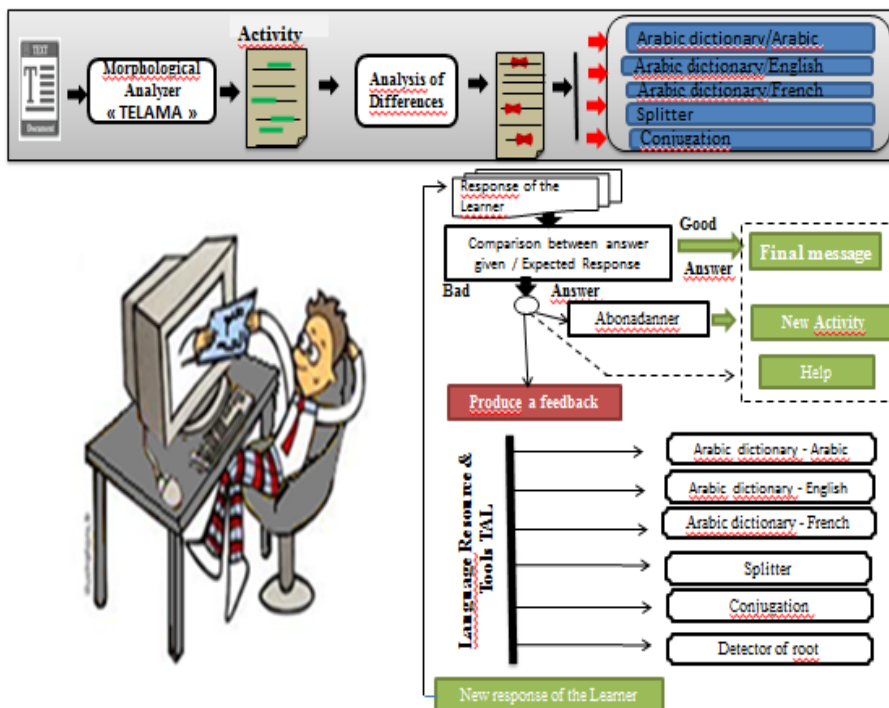
The teaching of language is based primarily on learning activities, which must be adaptable to the cultural level of the learner. To ensure a good learning approach, it is necessary to consider the activities to propose: the types of activities, the different levels of learning, the manner of presenting these activities (structures), the level of difficulty which is the most interesting for the allocation of aid, which depends on the type of the activity as well as the type of the error after the assessment [Antoniadis et al, 2006]. Teachers can describe activities, based on scripts describing exercises (for example, text to whole) and associating it with these exercises the functions NLP needed to build the activity (labelling morph syntactical and lemmatization in the case of text to hole for example). The creation of the activities can be arrange according of the scenarios created also by the teacher.

The system of language learning and especially the Arabic language is similar to the Artificial Intelligence, where the system must be intelligent to differentiate the type of assistance to assign to each type of activity using a multi system agent [1]. In the case of an explicit approach, i.e. insisting on the learning of the forms of the language, the NLP has a role of expertise on the productions of the learner, issuing the errors of the latter from a grammatical and lexical point of view. This vision is realized through the approach "tool" of the NLP.

Model of the Learner

The model of the Learner realizes, namely not only its content but also the diagnostiqueur, which is the set of processes that develop the model of the learner. It connects to the system as a simple visitor for his first time, using methods of access to the online sites like for example a search engine. It would therefore be invited to create a account, registering by the filling of a set of personal information, which, according to their nature may be mandatory or not, as essentially a login and a password that will be used by the following to save her answers. After be registered in the site, the latter becomes a number taking advantage of this system. It could therefore choose the level of difficulty and to respond to the questions of activities: each learner may choose learning activities or of assessment tests, corresponding to its level (a manual to aid exists in case of need).

The exercises that we propose are contextual exercises and other non-contextual who are not interested in the general context of the answers, but surrounded potentially of lexicons (spelling and phonemic, the syllabation, the grammatical category, the type and number, frequencies and lemmas associated, etc.). This choice of type of exercises is not arbitrary, but verified by the tools NLP which facilitate the management of this kind of activities. In

particular, the automatic generation of activities allows to focus activities on well-defined objectives (lexical language) and locate the media assistance, when we note that the knowledge of the learner are insufficient to do exercises for, advance to another level, make these activities interchangeable and to repeat, in New tests without the intervention of a teacher. The assessment of the responses should compare the response of the learner to the expected answer.

We have noted in particular that the individualization of learning remains the role the more assigned to the computer in the environments of the CALL. One of the key factors in an environment CALL, is the perception of the degree of interaction by the learners, who seems to have an impact on the learning and their satisfaction. It passes by the provision of learners of language resources own computerized, to provide them with the tools NLP of aid and by the articulation of these means around a pedagogical program relevant.

Statistics and Monitoring of Learning

Before the end of the development cycle of the environment TELA, we performed a few experiments in doing test the tool by a teacher and a learner. This was on the one hand to check if TELA brings to the teachers and learners most, for which it was originally designed, and on the other hand, to take into account the possible remarks, during development. The proposed environment provides a semi-automation appreciated by teachers, improvements on the ergonomic aspects which must be made.

The traces of registration are actions on which the learner can make a synthesis of the use of the system. The traces are organized by working session to save any action performed by the learner in date which allows to calculate the time spent in each response, in all

activities and a saw for the teacher on la way is used the system and to assess if, for example, an activity must be amended.

Table 1. Table of assessment of an activity

Time	Learner1	Learner2	Learner3	Learner4	Learner5	Learner6	Learner7	Average
Response 1	1	0.6	0.7	0.5	1	0.4	0.7	0.8222
Response 2	1	0.5	0.4	0.6	1	0.4	0.4	0.6666
Response 3	1	0.6	0.8	0.5	1	0.6	1	0.5888
Response 4	1	0.4	0.6	0.4	1	0.8	0.8	0.6555
Response 5	1	0.5	0.8	1	1	0.4	1	9
Response 6	1	0.6	0.5	0.6	1	0.5	0.7	0,8222
Response 7	0.9	0.4	0.6	1	1	0.7	0.4	0.5333
Response 8	1	0.	0.5	0.8	1	0.6	1	0,9444
Response 9	1	1	1	1	1	1	1	1
Response 10	1	0.6	0.6	1	1	0.5	0.8	0,7777
Activity	8.4	6.4	7.7	8.4	10	9.5	6.3	0.7444

The assessment of an activity can be carried out according to several parameters (number of correct answers, with the time to respond, the appreciation of the learner ...). In our example, the assessment is made according to the time spent by learners to take responses. It is therefore preferable to use methods of automatic processing of languages, which examine the productions of

learners with more flexibility, which allows a diagnosis more detailed and more reliable.

The Table 1, allows the teacher to conduct a comparative study at the level of the capacity of the understanding of each learner. Therefore, for our example, we note that the learner 2 is the best and the learner 5 is the lowest in this activity. This graph also shows that this activity is not reached, because we can see that the rate of success is more and more low at the end of the activity, which is generally opposed to the objective of the activity (more success toward the end of the activity) and shows that the purpose educational and teaching of this activity is not reached.

The recovery of traces on TELA allows the teacher to assess the learner. The tool retrieves any useful action carried out by the learner: The activities carried out, the time put to respond in each activity the scores obtained, and the number of correct answers. The teacher may as well proceed to a comparative study, at the level of the capabilities of understanding of each of the learners and to make a synthesis.

Conclusion

We have presented in this article a computing environment, favoring the acquisition lexical and the aid to the learning of the Arabic language based on tools of NLP with objectives and unlimited vocabularies and uses them as simple as possible.

This goal is accomplished through resources, methods and effective approaches in the different parts of our research work. This willingness to work is then the point of departure for new opportunities, to the intersection of two fields of research, NLP and CALL, which remained a long time separated. In this research work, we do a wide use at the crossroads of three disciplines: teaching of languages, NLP and informatics.

References

Alhoti, S. (2010). *Tag formalism for Arabic in view of educational applications and web implementation* (Doctoral dissertation). University of Franche-Comté, France.

Antoniadis, G., Echinard, S., Kraif, O., Lebarbé, T., Loiseau, M., & Ponton, C. (2004). NLP-based scripting for call activities: eLearning for linguistics and computational linguistics for eLearning. In *Proceedings of the International Workshop in association with COLING* (pp. 1–5). Geneva, Switzerland.

Bacha, K. (2018). Help system for creating educational resources for Arabic. *International Journal of Knowledge Systems Science, 9*(3), 32–47. https://doi.org/[Insert DOI if available]

Bacha, K., Zrigui, M., Mohsen, M., & Anis, Z. (2011). TELA: Toward environmental learning Arabic. In *Proceedings of the International Conference on Artificial Intelligence (CIIA'11)* (pp. 1–6). WORLDCOMP'11, Las Vegas, NV. Retrieved from http://cerc.wvu.edu/download/WORLDCOMP%2711/2011%20CD%20papers/EEE4685.pdf

Bacha, K., & Zrigui, M. (2012). Design of a synthesizer and a semantic analyzer's multi-Arabic, for use in computer-assisted teaching. *International Journal of Information Sciences and Application, 4*(2), 11–33.

Bacha, K., & Zrigui, M. (2012). Designing a model of Arabic derivation for use in computer-assisted teaching. In *Proceedings of the International Conference on Knowledge Engineering and Ontology Development (KEOD)* (pp. 1–5). Barcelona, Spain.

, K., & Zrigui, M. (2012). Machine translation system on the pair of Arabic/English. In *Proceedings of the International Conference on Knowledge Engineering and Ontology Development (KEOD)* (pp. 1–5). Barcelona, Spain.

Bacha, K., & Zrigui, M. (2014). Toward a model of statistical machine translation Arabic-French. In *Proceedings of the International Conference on Advanced Learning Technologies and Education (ICALTE)* (pp. 1–5). Hammamet, Tunisia.

Bacha, K., & Zrigui, M. (2014). Design and implementation of a model of segmentation environmental computer-assisted learning (TELA). In *Proceedings of the CIIA* (pp. 1–5). Ousse, Tunisia.

Bacha, K., Acheref, M., & Zrigui, M. (2014). Designing a model combination of Arabic for use in computer-assisted teaching. In *Proceedings of the World Congress on Computer Applications and Information Systems* (pp. 1–7).

Bacha, K., & Zrigui, M. (2015). Morphological analysis in the environment "TELA". *Procedia Computer Science, 62*, 191–215. https://doi.org/[Insert DOI if available]

Bacha, K., Jemni, M., & Zrigui, M. (2016). Toward a learning system based on Arabic NLP tools. *International Journal of Information Retrieval Research, 6*(4), 1–15. https://doi.org/[Insert DOI if available]

Bacha, K., & Zrigui, M. (2016). Contribution to the achievement of a spellchecker for Arabic. In *Proceedings of the 17th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing)* (pp. 1–5). Konya, Turkey.

Ben Othmane Zribi, C., Torjmen, A., & Ben Ahed, M. (2007). A multi-agent system for POS-tagging vocalized Arabic texts. *The International Arab Journal of Information Technology, 4*(4), 1–10.

Ben Mohamed, M. (2016). *Intégration des techniques de TAL dans l'apprentissage assisté par ordinateur pour l'enseignement de la langue arabe* (Doctoral dissertation). University of Sfax, Tunisia.

Bogaards, P. (1995). Dictionaries and written comprehension. *Cahiers de Lexicology, 67*(2), 37–53.

Bouzoubaa, K., & Souteh, Y. (2011). Vers une plateforme intégrée pour le traitement automatique de la langue Arabe. In *Proceedings of the TICLA Conference* (pp. 1–5). Paris, France.

Briot, P., & Demazeau, Y. (2001). Hermes models and architectures agents. In O. Boissier (Ed.), *Multi-Agent Systems* (pp. 1–5). Hermes.

Dichy, J., Braham, A., Ghazali, S., & Hassoun, M. (2002). The basis of linguistic knowledge DIINAR.1: Computerized dictionary of the Arab. In *Proceedings of the International Symposium* (pp. 1–5). University of Manouba, Tunisia.

Dichy, J., Hassoun, M., Mouelhi, Z., & Zaafrani, R. (2002). Toward a dictionary of Arabic (I): The words tools. *Linguistic Models, 36*, 72–88.

Dichy, J., & Farghaly, A. (2003). Roots & patterns vs. stems more grammar-lexis specifications: On what basis should a multilingual lexical database centered on Arabic be built? *Linguistic Studies, 45*(2), 1–10.

Bacha, K. (2019). Contribution to a new approach to analyzing Arabic words. In *Proceedings of ICCCI (2)* (pp. 46–57).

Cheikh, M., & Zrigui, M. (2020). Active learning-based framework for image captioning corpus creation. In *Proceedings of LION 2020* (pp. 128–142).

Khemakhem, A., Gargouri, B., Ben Hamadou, A., & Francopoulo, G. (2015). ISO standard modeling of a large Arabic dictionary. *Journal of Natural Language Engineering, 21*(4), 1–10. https://doi.org/[Insert DOI if available]

Jaballi, S., Manar Hazar, J., Nicolas, H., & Zrigui, M. (2023). Deep bidirectional LSTM network learning-based sentiment analysis for Tunisian dialectical Facebook content during the spread of the Coronavirus pandemic. In *Proceedings of ICCCI 2023 (CCIS Volume)* (pp. 96–109).

Lazhar, M., & Zrigui, M. (2022). Speaker naming in Arabic TV programs. *International Arab Journal of Information Technology, 19*(6), 843–853.

Shaalan, F. K. (2005). Intelligent computer-assisted language learning system for Arabic learners. *Computer Assisted Language Learning, 18*(1–2), 81–108.

Shaalan, F. K., & Talhami, E. H. (2006). Arabic error feedback in an online Arabic learning system. *Advances in Natural Language Processing, Research in Computing Science, 18*, 203–212.

Roman, A. (1997). *Grammar of Arabic*. Paris: Presses Universitaires de France.

Alhoti, S. (2010). *Tag formalism for Arabic in view of educational applications and web implementation* (Doctoral dissertation). University of Franche-Comté, France.

Antoniadis, G., Echinard, S., Kraif, O., Lebarbé, T., Loiseau, M., & Ponton, C. (2004). NLP-based scripting for call activities: eLearning for linguistics and computational linguistics for eLearning. In *Proceedings of the International Workshop in association with COLING* (pp. 1–5). Geneva, Switzerland.

Bacha, K. (2018). Help system for creating educational resources for Arabic. *International Journal of Knowledge Systems Science, 9*(3), 32–47. https://doi.org/[Insert DOI if available]

Bacha, K., Zrigui, M., Mohsen, M., & Anis, Z. (2011). TELA: Toward environmental learning Arabic. In *Proceedings of the*

International Conference on Artificial Intelligence (CIIA'11) (pp. 1–6). WORLDCOMP'11, Las Vegas, NV. Retrieved from http://cerc.wvu.edu/download/WORLDCOMP%2711/2011%20CD%20papers/EEE4685.pdf

Bacha, K., & Zrigui, M. (2012). Design of a synthesizer and a semantic analyzer's multi-Arabic, for use in computer-assisted teaching. *International Journal of Information Sciences and Application, 4*(2), 11–33.

Bacha, K., & Zrigui, M. (2012). Designing a model of Arabic derivation for use in computer-assisted teaching. In *Proceedings of the International Conference on Knowledge Engineering and Ontology Development (KEOD)* (pp. 1–5). Barcelona, Spain.

Bacha, K., & Zrigui, M. (2012). Machine translation system on the pair of Arabic/English. In *Proceedings of the International Conference on Knowledge Engineering and Ontology Development (KEOD)* (pp. 1–5). Barcelona, Spain.

Bacha, K., & Zrigui, M. (2014). Toward a model of statistical machine translation Arabic-French. In *Proceedings of the International Conference on Advanced Learning Technologies and Education (ICALTE)* (pp. 1–5). Hammamet, Tunisia.

Bacha, K., & Zrigui, M. (2014). Design and implementation of a model of segmentation environmental computer-assisted learning (TELA). In *Proceedings of the CIIA* (pp. 1–5). Ousse, Tunisia.

Bacha, K., Acheref, M., & Zrigui, M. (2014). Designing a model combination of Arabic for use in computer-assisted teaching. In *Proceedings of the World Congress on Computer Applications and Information Systems* (pp. 1–7).

Bacha, K., & Zrigui, M. (2015). Morphological analysis in the environment "TELA". *Procedia Computer Science, 62*, 191–215. https://doi.org/[Insert DOI if available]

Bacha, K., Jemni, M., & Zrigui, M. (2016). Toward a learning system based on Arabic NLP tools. *International Journal of Information Retrieval Research, 6*(4), 1–15. https://doi.org/[Insert DOI if available]

Bacha, K., & Zrigui, M. (2016). Contribution to the achievement of a spellchecker for Arabic. In *Proceedings of the 17th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing)* (pp. 1–5). Konya, Turkey.

Ben Othmane Zribi, C., Torjmen, A., & Ben Ahed, M. (2007). A multi-agent system for POS-tagging vocalized Arabic texts. *The International Arab Journal of Information Technology, 4*(4), 1–10.

Ben Mohamed, M. (2016). *Intégration des techniques de TAL dans l'apprentissage assisté par ordinateur pour l'enseignement de la langue arabe* (Doctoral dissertation). University of Sfax, Tunisia.

Bogaards, P. (1995). Dictionaries and written comprehension. *Cahiers de Lexicology, 67*(2), 37–53.

Bouzoubaa, K., & Souteh, Y. (2011). Vers une plateforme intégrée pour le traitement automatique de la langue Arabe. In *Proceedings of the TICLA Conference* (pp. 1–5). Paris, France.

Briot, P., & Demazeau, Y. (2001). Hermes models and architectures agents. In O. Boissier (Ed.), *Multi-Agent Systems* (pp. 1–5). Hermes.

Dichy, J., Braham, A., Ghazali, S., & Hassoun, M. (2002). The basis of linguistic knowledge DIINAR.1: Computerized dictionary of the Arab. In *Proceedings of the International Symposium* (pp. 1–5). University of Manouba, Tunisia.

Dichy, J., Hassoun, M., Mouelhi, Z., & Zaafrani, R. (2002). Toward a dictionary of Arabic (I): The words tools. *Linguistic Models, 36*, 72–88.

Dichy, J., & Farghaly, A. (2003). Roots & patterns vs. stems more grammar-lexis specifications: On what basis should a multilingual lexical database centered on Arabic be built? *Linguistic Studies, 45*(2), 1–10.

Bacha, K. (2019). Contribution to a new approach to analyzing Arabic words. In *Proceedings of ICCCI (2)* (pp. 46–57).

Cheikh, M., & Zrigui, M. (2020). Active learning-based framework for image captioning corpus creation. In *Proceedings of LION 2020* (pp. 128–142).

Khemakhem, A., Gargouri, B., Ben Hamadou, A., & Francopoulo, G. (2015). ISO standard modeling of a large Arabic dictionary. *Journal of Natural Language Engineering, 21*(4), 1–10. https://doi.org/[Insert DOI if available]

Jaballi, S., Manar Hazar, J., Nicolas, H., & Zrigui, M. (2023). Deep bidirectional LSTM network learning-based sentiment analysis for Tunisian dialectical Facebook content during the spread of the Coronavirus pandemic. In *Proceedings of ICCCI 2023 (CCIS Volume)* (pp. 96–109).

Lazhar, M., & Zrigui, M. (2022). Speaker naming in Arabic TV programs. *International Arab Journal of Information Technology, 19*(6), 843–853.

Shaalan, F. K. (2005). Intelligent computer-assisted language learning system for Arabic learners. *Computer Assisted Language Learning, 18*(1–2), 81–108.

Shaalan, F. K., & Talhami, E. H. (2006). Arabic error feedback in an online Arabic learning system. *Advances in Natural Language Processing, Research in Computing Science, 18*, 203–212.

Roman, A. (1997). *Grammar of Arabic*. Paris: Presses Universitaires de France.

Acknowledgements: