Towards a Learning System Based on Arabic NLP Tools

Khaireddine Bacha, LaTICE Laboratory, University of Tunis, Tunisia
Mohamed Jemni, LaTICE Laboratory, University of Tunis, Tunisia
Mounir Zrigui, LaTICE Laboratory, University of Tunis, Tunisia

ABSTRACT
The evolution of computer technology stimulates us to develop a learning environment that addresses many of the problems associated with the use of words in Arabic. A learning platform of Arab foreign language used automatic language processing tools (NLP) based on multi-agent approach. Initially, this environment gives teachers the ability to automatically design other variants of educational activities, which notably includes automatic evaluations. Then it enables the learner to build his or her language skills and to learn concepts of the Arabic language.

KEYWORDS
Arabic Learning, Educational Activities, Environment, Learner, Multi-agent, Teachers, Tools (NLP)

1. INTRODUCTION

The Information Technology and Communication (ICT), namely the Internet, have, recently, been invading our daily lives at both levels: personal and professional. Having interfered in many areas such as traditional commerce and government (e-government), Internet is indeed, becoming the cornerstone of a new form of education. In fact, the multiple online learning sites bring gains in terms of time and flexibility in addition to significant interactivity.

Some learning arrangements of Arabic exist and are endowed with some features in an attempt to meet the specific needs of this type of training (monitoring learners, collaborative work management system, etc.). It, therefore, seemed appropriate to provide us with other devices and tools and automate benefit processing of language in these platforms.

Virtually automatic processing language (TAL) capable of implementing learning activities, linguistic information and provides adaptable pedagogy return trick with students the skills. The use of NLP, for designing CALL software, is a reliable way to solve these problems (Nielsen & Carlsen, 2003; Gheith, Dawa, & Afifty, 1996).

Despite the development of the Arab TAL, in recent years, the number of tools that are available to the CALL to provide teachers and learners with interactive environments for learning the Arabic language remains, in our view, modest. Yet, some achievements have emerged. Regardless of the techniques used and the value they have added these learning systems are still quite rare. Actually, all of them attempted to change the computer assisted learning. Some achievements in this direction were reached such as conventional tools like hotpotatos, Netquiz, etc., along with other tools using NLP like:

• “ArabVISL” or also “Arabic CALL”. This system allows learners to produce sentences in different contexts. According to Shaalan, the system allows learners to recognize mistakes and
guide them to a self-correction (Shaalan & Kaled, 2005; Shaalan & Talhami, 2006; Verlinde, Selva, & Binon, 2003).

- One of the first tools is made for the detection of root words (verbal excesses, nominal derivatives) (Mars, Antoniadis, & Zrigui, 2008). This system was then adapted to detect and correct learners’ errors.
- “SALA” (Nielsen & Carlsen, 2003), this tool provides learners with Arabic derivations exercises according to patterns or a combination of activities. It uses a conjugator and a noun and verb derivation tool. The Feedback is still classic the True / False style.
- A final learning tool based on NLP tools, called @rab-Learn was, also, developed (Mars & Antoniadis, 2012). It is a learning tool, including solutions from the TAL. It is easy to use and it allows teachers of Arabic, non-specialist NLP, to use the resources as well as to generate the set of activities for learning. It is therefore unnecessary to use a chain of complex treatments, or to make use of data processing specialists or TAL (Maraoui, Antoniadis, & Zrigui, 2009).

However, they both have common boundaries in terms of variety of activities they offer, interaction, openness and quality feedback. These are its limits:

- For conventional tools presented, they often look like traditional exercises that can be found in the books from which they are inspired.
- The feedback in these tools do not adjust to learners’ level or nature of activities. Corrections are preset (the system compares the response of the learner with the predefined) and offer no explanation as to the nature of the error.

All these tools have common boundaries in the absence of feedback (or automatic correction or assistance if needed), rigidity because the data used are often static and limitations of numbers of learners covered since the exercises are usually in Arabic (used only from Arabic).

It is in this context that lies our research. Hence, we will introduce a structured learning environment that takes account of the use of automatic language processing tools (TAL) to learn Arabic. These linguistic resources will be used for the generation of educational activities for the teaching of Arabic language. Using the result for educational applications under the TELA environment: This is supposed to be a platform dedicated to language teachers which allows them to create educational activities in their area of interest and offers exercises on different properties words, immediate correction, feedback (feedback-based agent-oriented architecture).

2. NLP INTEGRATION LEVEL IN TELA

The overall architecture of the learning environment TELA “Towards Environmental Learning Arabic” (Bacha, Zrigui, Mohsen, & Anis, 2011), allowing language teachers to create learning activities based on technologies from the TAL, which notably includes automatic assessments. This environment was developed according to a principle of modular design for maximum flexibility and scalability. Needless to say that this approach was not an optional choice but rather a necessity justified by two points. The first point is that the system is too complex to be represented by a single module. The second point is that the system is subject to possible extensions that can hit several levels since the acquisition of the texts and the display of results.

We present in this part, one of the key steps in achieving a learning environment of Arabic; it is initially to segment an Arabic text into paragraphs, sentences or words relying on the strategy based on finite state automata (Bacha, Acheref, & Zrigui, 2014), and then the labeling phase, is performed using a linguistic development environment that allows linguists to formalize various levels of linguistic phenomena and able to provide different interpretations of each analyzed form. Then, the
Pharaoh: Context-Based Structural Retrieval of Cognitive Scripts
www.igi-global.com/article/pharaoh-context-based-structural-retrieval/78314?camid=4v1a