Gamification for Arabic Natural Language Processing: Ideas into Practice

Imad Zeroual, Anoual El Kah and Abdelhak Lakhouaja
Computer Sciences Laboratory, Faculty of Sciences, Mohammed First University
Oujda, Morocco
mr.imadine@gmail.com, elkah.anoual.mri@gmail.com, abdel.lakh@gmail.com

ABSTRACT

Gamification is a novel IT development that focuses on how to address the use of games and apply game design techniques in a non-gaming context. Recently, it has significantly attracted several researchers' attention concerned with different fields such as education, business, and information retrieval, among other fields. Most Natural Language Processing (NLP) applications use large accurately labelled data sets to achieve good performance, but these data sets are hard to obtain. The standard method to produce labelled data has been through manual labour that demands a strong engagement and a heavy financial participation. For these reasons, gamification may offer smooth possibilities for wide improvements in this field. However, gamification by itself may not yield the widely-expected results if the incentives are not strong motivators. Further, based on a previous experience, we strongly believe that even financial incentives are not enough for the participants to increase their contributions. Therefore, we could incentivise the participants, through gamification, and through their need for language acquisition. Thus, language-based games are adopted in this project as a direct incentive and a stronger engagement to ensure the benefits for both NLP practitioners and the language learners. In this paper, we present the first Arabic project that targets all the means to promote and explore the possibilities for research and practical applications of using games and gamification for Arabic NLP. Here, we shed light over the advantages of using a gamified approach in corpus annotation, named entity recognition, and word sense disambiguation.

Keywords—Gamification, Games design elements, Natural language processing, language-based games, Arabic language.

1   Introduction

Basically, gamification originates from the computer games industry. It is based on applying game thinking and game mechanics or game design techniques to non-game activities in order to engage users, solve problems, and drive behaviour [1]. Gamification is considered to be an effective approach to increase users' motivations in non-game activities or towards using technology; thereby, increasing the quantity and quality of the output of the given activities [2].

Since 2015, millennials became the largest generation in the workforce [3]; yet, the world's population spends roughly 3 billion hours a week playing video games [4]. Thus, gamification is becoming increasingly

important especially with the tremendous growth of the smartphone and the increased use of the Internet. This raises the gamification implementations in education [5], information retrieval [6], healthcare [7], business [8], tourism [9], and insurance markets [10], among other fields.

On the other hand, most Natural Language Processing (NLP) applications require a strong engagement and greater participation which is often not available to reach targeted purposes or solve some issues, primarily for morphologically complex and resource-poor languages such as Arabic. Due to these convincing reasons, we believe that gamification and language-based games could incentivise the learners to continue for their contributions and ensure the benefits for the NLP practitioners. For example, games can be used to collect large numbers of annotations provided by sufficient numbers of learners who are motivated to play. Thereby, the aggregated annotations can be used as labels that complement a previous automatic annotation and may replace the post-editing by a domain expert, which is often difficult to have someone available, at least for the Arabic language.

In this paper, we present the first Arabic project that aims promoting and exploring the possibilities for research and practical applications of using gamification for Arabic NLP. The main objective is to provide recommendations collected from existing research and to discuss how the Arabic NLP researchers and practitioners can benefit from using games and gamification. Further, a preliminary approach for developing and testing some proposed language-based games is presented. This gamified approach is primarily used in corpus annotation, named entity recognition, and word sense disambiguation.

The paper is arranged in five main sections focusing on the basic aspects of gamification and the possibilities of using it for Arabic NLP. In section 2, a background information regarding Arabic NLP is provided. In section 3, we briefly describe the standard basics of gamification and the known outcomes of its implementation. Further, guidelines and recommendations for developing games primarily language-based games are presented. In section 4, the proposed gamified approach and a set of preliminary games are described. Finally, we conclude this paper in section 5.

2 Arabic Natural Language Processing

Arabic is currently the most used Semitic language [11]. It is expanding in the world in an area extending from the Arabian/Persian Gulf in the East to the Atlantic Ocean in the West. According to UNESCO, the Arabic language is spoken by more than 422 million native speakers [12] around 29 countries and among 1.6 billion Muslims worldwide use it to perform their daily prayers [13]. Moreover, the presence of the Arabic language on the internet grew around 6.091% in the last fifteen years (2000–2015), it is the highest growth of the ten top online languages, and it is the fourth most spoken language on the internet [14]. Adding the fact that teaching Arabic as a foreign language has become a global educational enterprise [15].

2.1 Arabic language features

The standardization of Arabic lexicon and grammar are deeply rooted and well established a long time ago. The morphology of Arabic differs in the structure of affixes from Indo-European languages [16] (e.g. English, German, French, Hindi). Arabic is a highly productive both derivationally and inflectionally [17]. These two complex paradigms are based on the interaction between roots and patterns which have intrigued lexicographers and morphologists for centuries.

Basically, Arabic language consists of three main categories [18], [19]: Noun "اسم" <Asm>, Verb "فعل" <fEl> and Particle "حرف" <Hrf>. In addition, each one of these categories has dozens of subcategories [20]. One of the most challenging and interesting property of Arabic is the concatenative morphemes, where one word could represent a whole sentence through sequential concatenation. For example: the Arabic word "أَسَأَلْتُمُونِيهَا" <AasaAalotumuwniyhaA> means in English "Do you have already asked me for it".

The omission of diacritics (short vowels) in written Arabic is another source of some difficulties in several automatic processing systems of Arabic language [21]. When these vowels are omitted it is left for the reader to infer; knowing that the vowels can encode grammatical category or feature information. Typically, a root or a lexical item consists of three consonants. The vowels add grammatical information when attached to these consonants. For example, the three consonants ktb can stand for the verb "كَتَبَ" <kataba> meaning "he wrote", or for the plural noun "كُتُبٌ" <kutub> "books".

Another characteristic of the Arabic language is the free word order nature in the sentence which makes parsing Arabic sentences one of the most difficult tasks. i.e., we could easily change the order between the subject and the verb; yet, it is not necessary that they must agree with one another in number (singular or plural) for all cases. For example:

1. The boy hit the ball

ضَرَبَ الْوَلَدُ الْكُرَةَ or الْوَلَدُ ضَرَبَ الْكُرَةَ

2. The children watch television

يُشَاهِدُ الأَطْفَالُ التِّلْفَازَ or الأَطْفَالُ يُشَاهِدُونَ التِّلْفَازَ

Although the free order of the subject "الْوَلَدُ" <Alowaladu> "The boy" and the verb "ضَرَبَ" <Daraba> "hit"; yet, the subject–verb number agreement of the subject "الأَطْفَالُ" <AloATofaAlu> "The Children" and the verb "يُشَاهِدُ/ يُشَاهِدُونَ" <yu$aAhidu/yu$aAhiduwna> "watches/watch" has changed, both sentences are grammatically correct and have the same meaning. And this becomes more complex in longer sentences.

2.2 Arabic text processing

In this section, we present some well-known tasks in Arabic NLP fields and which will be under consideration for the gamification implementation.

2.2.1 Corpus annotation

One of the main aims being addressed in corpus linguistics is to develop new forms of annotation and improve the accuracy of automatic annotation. The following tasks are some of the key features of text annotation:

- Lemmatization: It is relating words written in different syntactic forms to their canonical base representation (i.e., lemma). The lemma is a dictionary lookup form which can relate different word forms that have the same meaning [22]. For instance, the lemma for the verbs "يَكْتُبُونَ" <yakotubwna> "they write" and "تَكْتُبُ" <takotubu> "she writes" is "كَتَبَ" <kataba> "to write".

- Part of speech tagging: It is a basic task in many fields primarily corpus linguistics and NLP. It aims to automatically assign a morpho-syntactical label (e.g., noun, verb, and adjective) to every word in a written text. Further, this task allows simple syntactic searches to be performed.
- Parsing: It is the natural successor to part of speech tagging. Basically, it provides a dependency tree as an output. The goal of parsing is to predict for each sentence or clause an abstract representation of the grammatical entities and liaisons between them. Consequently, the parser assigns a fully labelled syntactic tree or bracketing of constituents to sentences of the corpus [23]. For example, Figure 1 exhibits a parse Tree of the first verse of chapter 113 (surah Al-Falaq) of the Holy Quran "قُلْ أَعُوذُ بِرَبِّ الْفَلَقِ" "Say, 'I seek refuge in the Lord of daybreak'" which is parsed using NLTK [24].
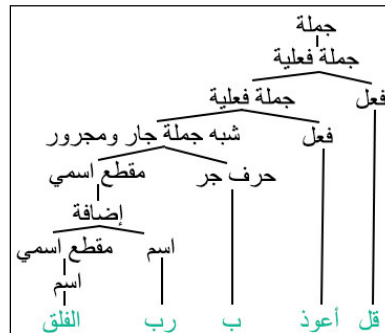

Figure 1 A parse Tree of the first verse of chapter 113 of the Holy Quran.

2.2.2 Named entity recognition

Named entity recognition (NER) is an information extraction technique that aims to locate and categorize proper nouns (e.g., names of persons, organizations, and locations) in a text [25]. NER plays a vital role in several NLP applications; it provides key information that are very useful for Question Answering and Machine Translation.

Work related to Arabic NER is rather limited and under-studied as a recent survey reported [26]. One of the main reasons that led to this situation is the absence of a large labelled corpus which is required to train NER models. In addition, the mentioned characteristics of the Arabic language make this task very challenging. Figure 2 shows a sample NER output for an Arabic sentence:

"يوم الجمعة الماضية، قام الملك محمد السادس بافتتاح مركز الدراسات والبحوث الانسانية والاجتماعية بمدينة وجدة."

"Last Friday, King Mohammed VI inaugurated the Center for Human and Social Studies and Researches in Oujda City."

يوم <DATE>الجمعة<DATE/> الماضية، قام الملك <PER/>محمد السادس<PER> بافتتاح <ORG/>مركز الدراسات والبحوث الانسانية والاجتماعية<ORG> بمدينة <LOC/>وجدة<LOC>.

Figure 2 A sample NER output for an Arabic sentence.

2.2.3 Word sense disambiguation

Word Sense Disambiguation (WSD) is a technique to determine automatically the correct or the appropriate sense of each ambiguous word regarding its context [27]. WSD is one of the most difficult

problems in NLP that negatively affects the accuracy of information retrieval systems due to the sparse nature of the queries involved. It exerts a similar influence upon machine translation [28]. It is ubiquitous across all languages but it has greater challenges in the Arabic language [29] primarily due to these two issues:

- A high level of ambiguity emerges when diacritics are omitted, especially in most Modern Standard Arabic (MSA) texts. For example, the unvowelled isolated word "كتب" <ktb> can mean the word "كَتَبَ" <kataba> "to write", "كُتِبَ" <kutiba> "written", and the word "كُتُب" <kutub> "books", among other meanings.

- Part of speech (PoS) category of a word reflects its syntactic and semantic role in a given context. What's more, it is estimated that the average number of possible PoS tags for a word in most languages is 2.3, whereas in MSA is 19.2 [30]. For instance, the same vowlled word "أَحْمَدُ" <AHomadu> has different meaning according to its PoS category as in the sentence "أَحْمَدُ اللَّه" "I thank Allah" it is a verb and as in "أَحْمَدُ بنُ عَبْدِ اللَّه" "Ahmed is Abdullah son's" it is a proper noun.

3 Gamification

Gamification is the use of game design elements in non-game contexts to increase users' motivations towards given activities [31]. Gamification was first mentioned in 2003 and became widely used in literature in 2010 [32]. Over these years, gamification has attracted a great deal of attention, from all shapes and sizes of research groups in academia and industry alike, in order to revolutionize the way they interact with targeted users and customers. Further, gamified approaches become more competitive and more profitable for data collection. It addresses the limitations of other methods such as the manual, automated, and crowdsourcing approaches [33]. In its core, gamified approaches are based on the same strategy as crowdsourcing. However, in this latter, the participants receive money to increase motivation; whereas, gamification incentivises them by getting entertained.

The major keys benefits of gamification are a better learning experience, instant feedback, and prompting behavioural change. Further, gamification techniques can provide comprehensive navigation and frequent indications to the user, which keep the user informed of what to do and explaining how answers are correct or incorrect [34]. In addition, in the areas of education and business, gamification can be used to fulfil most learning needs including induction and onboarding, product sales, customer support, soft skills, awareness creation, and compliance. Using wise strategy and high levels of engagement, gamification leads to an increase in recall and retention. It can drive strong behavioural change, especially when combined with the scientific principles of repeated retrieval and spaced repetition.

Several studies have integrated gamification concepts and benefits in the NLP and language learning basic skills. A previous study was conducted to design and develop an accessible web application, named Playlingua, to improve the Spanish language learning process and lecture comprehension through gamification [35]. Another study aimed to build gold standard data for word sense disambiguation using a "Game with a Purpose" (GWAP) called Wordrobe, This game consists of a large set of multiple-choice questions on word senses generated from the Groningen Meaning Bank [36].

The coming section is devoted to providing guidelines and recommendations for developing games and language-based games.

3.1 Games and their elements in Gamification

Gamification is about more than just playing games. It is defined as the concept of applying game-design thinking to non-game applications. It has been identified as a promising technique to improve students' engagement and encourage desired behaviours which could have a positive impact on learning in general and language learning in particular.

Games play a vital role in enhancing the children's knowledge and attitudes. It presents an opportunity to acquire knowledge and expose to both failure and success experiences, accordingly, the child gains good morals and attitudes [37]. During the game experience, children are motivated and interact in multi-sensory ways while enhancing their skills and renewing their motivation in every replay [38]. Therefore, more interest is given to edutainment and pedagogical-based games which bind knowledge with entertainment.

Viewing that designing games is a very multiplex activity; it is recommended that game design should be a team process. For this reason, game development teams should include interfaces specialists, game designers, as well as subject matter experts, cognitive task analysis, and game research in case consultation is needed [39].

Many effective techniques and procedures for designing games are set for the reason to design effective games that reliably reach intended objectives, as it is mentioned below [40]:

✓ Interface design is preferable to be colourful to grab the children's attention in the fact that children lose quickly their interest primarily toward educational tasks.

✓ An instant feedback after every activity or action should be provided for the reason to inform the child whether the answer was correct or needs another try.

✓ Wisely, clear sounds effects and instructions with suitable speed should be used to enrich the interactive experience of children.

✓ Icon should be accompanied with a label in order to allow children, regardless of their cognitive disability, to understand what expected from them when clicking on that icon.

✓ The set of games should be mapped in an understandable structure, especially if it contains multiple levels.

✓ The physical effort should be deduced, in a way that the game can be used efficiently and comfortably by either normal children or those with fine motor skills. This aim is fulfilled by using easy drag and drop functionalities and easy target selection facility.

According to pioneer studies [31], [41], [42] in the gamification field, the success of gamified systems will increase if the following game elements are considered in the design as much as possible:

✓ *Self-representation with avatar or virtual guidance;*

✓ *Points, scores, if they are exchanged for badges or rewards, they can become more effective;*

✓ *Leaderboards;*

✓ *Feedback;*

✓ *Classified Levels; to unlock new levels by finishing previous ones increases the self-confidence of the player;*

✓ *Story/Narrative context;*

✓ *Exercises (Challenges, Discoveries etc.);*

✓ *Time pressure (Countdown, Speed etc.);*

In addition to the previous elements, Morschheuser et al. [2] recently investigates overall instructions proposed by 25 experts from 16 different countries that have real-world gamification project experience. He claims that identifying the problems that should be addressed via gamification and clarifying the project's objectives and how they can be measured is the main criterion in gamified applications and systems.

3.2 Language-based games

There is an increasing use of games in learning contexts. They come in different technologies and are implemented in diversified educational settings. Well-designed games might have great potential for optimizing the learning achievements of the children, additionally, they can enhance their learning skills independently on time or places where they are, and while travelling, shopping or simply when being at home. The use of games for language learning is labeled game-based language learning, these games are often associated with the approach of task-based language teaching focusing on meaningful language use rather than drill-and-practice exercises [43]. Edutainment games, primarily language-based games, should take into consideration elements of engagement, motivation, and enjoyment. These games must be intuitive and use less cognitive load in order to ensure that the learning experience is essentially more engaging and worthwhile.

Despite the growing popularity and the promising claims that have arisen, there seems to be a major lack of empirical research and evidence on the added value of game-based learning [43], whilst, other several studies explored gaming for a specific purpose in language learning, such as listening strategies, and English as a foreign language, writing and speaking performance using a multimedia web annotation system [44].

The great challenge of developing educational games is to keep a balance between learning and gaming, and between challenge and users' abilities [45]; besides, integrating the motivating aspects of games with a good instructional design is critical. Nevertheless, effective techniques and procedures for designing games that reliably reach intended objectives are still not enough and do not fit a wide range of users including those with disabilities.

Unfortunately, the learning resources for Arab children, primarily Edutainment games, are very limited. Furthermore, the developed games and apps are not compatible with the culture, language, and background of the Arab children. In fact, less than 1% of available apps are Arabic-based [46]. Consequently, this project contributes to increase the Arabic resources by developing a suitable set of games for Arab children; yet, serving the Arabic NLP.

4 The proposed gamified approch and games

The main motivations of the proposed gamified approach are amusement, getting "knowledge", and learning a language. Thereby, learners will use the set of games for fun and learning purposes. On the other hand, the act of collecting their responses in the games is essential since it provides valuable contributing data for Arabic NLP and keeps traces regarding learner's status and comprehension of the game [39].

4.1 Gamified approach

The current work is a collection of games with purposes; each game targets a specific task of Arabic NLP. The proposed gamified approach is primarily used for corpus annotation, named entity recognition, and word sense disambiguation. The proposed set of games attempts to respect the recommended techniques and procedures for designing games in order to reach the intended objectives of our approach. It provides an instant feedback, a simplified pedagogical content, ensured engagement, well-designed interfaces, and last but not least, interaction that grabs the learners' attention and increases their motivation towards language learning.

The developed games are exploited to collect valuable data from the player's interactions. For example, games can collect large numbers of annotations information provided, since sufficient numbers of players are motivated to play. After the aggregation of the annotations, they can be used to replace or compliment the effort of NLP practitioners, primarily expert annotators. The games are designed for non-expert users, who will use their intuitions about language to annotate linguistic phenomena without being disturbed by technical terminology.

However, intensive research and studies should investigate the criteria followed to obtain reliable and accurate results such as:

- Participant's number: investigate which number of the participants is sufficient to affirm that the annotation is correct;
- Learners' score: discuss the achieved score which ensures that the player is able to annotate the raw data;

Level difficulty: evaluate the impact of the level difficulty on the performance of the player in the annotating results.

Besides, an evaluation of the games should be led to evaluate both the games' content difficulty and the player's performance.

4.2 Technical aspect

Recently, an emerging technology has led to focus more on cross-platform applications instead of device dependent application. This trending concept facilitates the access independently of the type of the device used. Multi-platform is a term for computing methods and programs that are inter-operating on different platforms and devices. Regarding the cross-platform advantages, we intentionally adopted the concept of building once and publishing everywhere.

The games are developed using a 2D game engine based on HTML5 named Construct2[1]. It creates games using a visual editor and a behaviour-based logic system. The exportation from this editor to most major platforms is allowed and the access from different devices is assured through its supported platforms: Android, Windows Phone, Windows Desktop, Mac Desktop, Linux Desktop…

---

[1] www.scirra.com

4.3   The proposed games

Basically, the following set of games contains three principal parts, each part targets respectively, PoS tagging, lemmatization, and parsing. These parts are the key features of text annotation.

The entire games share to some extent the same strategy: multiple-choices which could be dragged and dropped to the right place. This method provides less kinaesthetic efforts which make the games more suitable for children with fine motor skills.

Failure or success sound effects are involved in the games in addition to the instructions which are recorded and recited by a virtual companion. We adopt the auditor strategy for the sake of reducing on-screen text which could be distracting for some users with reading deficits. The oral instruction may enhance the listening comprehension of the player.

Players are restricted by timing that varies according to the game level (From 20 to 60 seconds); the harder the level is the more time is given. Players are rewarded with golden stars collected all over the set of games.

As mentioned, the set of games is beneficial for both NLP practitioners and the language learners. On one hand, the games are edutainment games that address learning Arabic language, on the other hand, we, as NLP practitioners, will benefit from the collected annotated data of the players' interactions. In the following, we describe the design of each game.

4.3.1   PoS tagging games

PoS tagging games consist of displaying a sentence divided into words. The player should define the correct PoS tag for each word in the given context by dragging the selected word to its correct morpho-syntactical label. **Error! Reference source not found.** presents an example of a screenshot of the game that focuses on nouns, verbs, and particles. It includes more categories in further levels such as adjectives and adverbs.

Figure 2 Screenshot of an example of PoS tagging games

4.3.2   Lemmatization games

Lemmatization games consist of binding words with their canonical base representation (i.e., lemma). **Error! Reference source not found.** demonstrates one of the lemmatization games; it is based on relating different syntactic forms to their lemma. The number of the suggested lemma in this game is restricted to four lemmas, it increases accordingly with which level of the game is played.

Figure 4 Screenshot of lemmatization game.

4.3.3   Parsing games

Parsing games aim to predict for a proposed sentence an abstract representation of the grammatical entities and liaisons between them. During these games, the player should assign syntactic labels to the component of the suggested sentence. **Error! Reference source not found.** shows a concrete example of this games' category.

Figure 5 Screenshot of a parsing game.

5 Conclusion and perspectives

The chief purpose of this project is using gamification for a relatively resource-poor language such as Arabic. In this paper, we present Arabic language features which differ from Indo-European languages. Further, we suggest the use of gamified approaches as competitive and alternative to produce labelled data for the Arabic NLP applications. Yet, they are more profitable for data collection and address the limitations of other methods such as the manual, automated, and crowdsourcing approaches.

Gamification improves participants' engagement by making their contribution to the task more efficient, the collection of data much faster, and the cost does not scale with the amount of data. As a case study, we developed a set of preliminary games. Although it is not expected for these games to match the quality of Arabic NLP experts, they are able to address the limitations of the other approaches and still provide good quality results. However, after the aggregation of the collected data, they can be used to replace or compliment the effort of NLP practitioners, primarily expert annotators.

Finally, this work is another step to demonstrate the advantages of using a gamified approach over other methods of data collection for Arabic language and to minimize the hand-correction. Later, intensive

research and studies should investigate the criteria followed to obtain reliable and accurate results such as participant's number, learners' score, and level difficulty. Besides, an evaluation of the games should be led to evaluate both the games' content difficulty and the player's performance.

REFERENCES

[1]   B. Kim, *Understanding gamification*. American Library Association, 2016.

[2]   B. Morschheuser *et al.*, "How to gamify? Development of a method for gamification," in *Proceedings of the 50th Annual Hawaii International Conference on System Sciences (HICSS)*, 2017, pp. 4–7.

[3]   L. Naik, "New liquid workforce your competitive advantage: feature-workplace planning," *HR Future*, vol. 2016, no. Oct 2016, pp. 26–27, 2016.

[4]   J. Young, "Can Library Research Be Fun? Using Games for Information Literacy Instruction in Higher Education," *Ga. Libr. Q.*, vol. 53, no. 3, p. 7, 2016.

[5]   B. Taspinar, W. Schmidt, and H. Schuhbauer, "Gamification in Education: A Board Game Approach to Knowledge Acquisition," *Procedia Comput. Sci.*, vol. 99, pp. 101–116, 2016.

[6]   C. I. Muntean and F. M. Nardini, "Gamification in Information Retrieval: State of the Art, Challenges and Opportunities.," in *IIR*, 2015.

[7]   [7]   D. King, F. Greaves, C. Exeter, and A. Darzi, *"Gamification": Influencing health behaviours with games*. SAGE Publications, 2013.

[8]   D. Gears and K. Braun, "Gamification in business: Designing motivating solutions to problem situations," in *Proceedings of the CHI 2013 Gamification Workshop*, 2013.

[9]   F. Xu, J. Weber, and D. Buhalis, "Gamification in tourism," in *Information and Communication Technologies in Tourism 2014*, Springer, 2013, pp. 525–537.

[10]   S. D. Ali, "Game Theory & Gamification Applications to Property & Casualty Ratemaking," 2016.

[11]   M. N. Al-Kabi, S. A. Kazakzeh, B. M. A. Ata, S. A. Al-Rababah, and I. M. Alsmadi, "A novel root based Arabic stemmer," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 27, no. 2, pp. 94–103, 2015.

[12]   I. Bokova, *World Arabic Language Day*. 2012.

[13]   M. B. Yassein and Y. A. Wahsheh, "HQTP v. 2: Holy Quran Transfer Protocol version 2," in *Computer Science and Information Technology (CSIT), 2016 7th International Conference on*, 2016, pp. 1–5.

[14]   S. Hadžiahmetović Jurida, M. Džanić, T. Pavlović, A. Jahić, and J. Hanić, "Netspeak: Linguistic Properties and Aspects of Online Communication in Postponed Time," *J. Foreign Lang. Teach. Appl. Linguist.*, vol. 3, no. 1, pp. 1–19, 2016.

[15]   M. L. Sakho, "Teaching Arabic as a Second Language in International School in Dubai A case study exploring new perspectives in learning materials design and development," British University in Dubai, 2012.

[16]   I. K. Gharaibeh and N. K. Gharaibeh, "Towards Arabic noun phrase extractor (ANPE) using information retrieval techniques," *Softw. Eng.*, vol. 2, no. 2, pp. 36–42, 2012.

[17]   N. Alsaedi, P. Burnap, and O. F. Rana, "Sensing real-world events using Arabic Twitter posts," 2016.

[18]   A. Al-Dahdah, "The grammar of the Arabic language in tables and lists," *Beirut Maktabat Lebnan Arab.*, 1989.

[19]   M. I. M. S. Ghalayini, *Jami'al-durus al-'arabiyah*. Turath For Solutions, 2013.

[20]   I. Zeroual, A. Lakhouaja, and R. Belahbib, "Towards a standard Part of Speech tagset for the Arabic language," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 29, no. 2, pp. 174–181, 2017.

[21]   A. Chennoufi and A. Mazroui, "Impact of morphological analysis and a large training corpus on the performances of Arabic diacritization," *Int. J. Speech Technol.*, vol. 19, no. 2, pp. 269–280, 2016.

[22]   M. Attia and J. Van Genabith, "A jellyfish dictionary for Arabic," in *Electronic lexicography in the 21st century: thinking outside the paper: proceedings of the eLex 2013 conference, 17-19 October 2013, Tallinn, Estonia*, 2013, pp. 195–212.

[23]   R. Tsarfaty, D. Seddah, S. Kübler, and J. Nivre, "Parsing morphologically rich languages: Introduction to the special issue," *Comput. Linguist.*, vol. 39, no. 1, pp. 15–22, 2013.

[24]   M. Shatnawi and B. Belkhouche, "Parse Trees of Arabic Sentences Using the Natural Language Toolkit," *Coll. IT UAE Univ. Al Ain*, 2012.

[25]   B. Mohit, "Named entity recognition," in *Natural Language Processing of Semitic Languages*, Springer, 2014, pp. 221–245.

[26]   A. Dandashi, J. Al Jaam, and S. Foufou, "Arabic Named Entity Recognition—A Survey and Analysis," in *Intelligent Interactive Multimedia Systems and Services 2016*, Springer, 2016, pp. 83–96.

[27]   A. R. Pal and D. Saha, "Word sense disambiguation: A survey," *ArXiv Prepr. ArXiv150801346*, 2015.

[28]   A. Zouaghi, L. Merhbene, and M. Zrigui, "A hybrid approach for arabic word sense disambiguation," *Int. J. Comput. Process. Lang.*, vol. 24, no. 02, pp. 133–151, 2012.

[29]   M. Alian, A. Awajan, and A. Al-Kouz, "Arabic Word Sense Disambiguation Using Wikipedia," *Int. J. Comput. Inf. Sci.*, vol. 12, no. 1, p. 61, 2016.

[30]   A. Farghaly and K. Shaalan, "Arabic natural language processing: Challenges and solutions," *ACM Trans. Asian Lang. Inf. Process. TALIP*, vol. 8, no. 4, p. 14, 2009.

[31]   S. Deterding, D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: defining gamification," in *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*, 2011, pp. 9–15.

[32]   L. Lund and P. O'Regan, *Gamifying Natural Language Acquisition: A quantitative study on Swedish antonyms while examining the effects of consensus driven rewards*. 2016.

[33]   T. J. Tiam-Lee and S. See, "Building a sentiment corpus using a gamified framework," in *Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), 2014 International Conference on*, 2014, pp. 1–8.

[34]   D. Barisas, A. Duracz, and W. Taha, "DSLs Should be Online Applications," *2014 Jt. Int. Conf. Eng. Educ. Int. Conf. Inf. Technol. 2-6 June 2014 Riga Latv.*, pp. 314–319, 2014.

[35]   P. Torró, E. and Lloret, E. and Moreda, "Playlingua: Exploring the Application of Natural Language Processing Technologies and Gamification in Language Learning," *Edulearn15*, no. April, 2015.

[36]   N. J. Venhuizen, V. Basile, K. Evang, and J. Bos, "Gamification for Word Sense Labeling," *Proc. 10th Int. Conf. Comput. Semant.*, pp. 397–403, 2013.

[37]   R. Skiada, E. Soroniati, A. Gardeli, and D. Zissis, "EasyLexia: A Mobile Application for Children with Learning Difficulties," *Procedia Comput. Sci.*, vol. 27, pp. 218–228, 2014.

[38]   C. Foscari and F. Caon, *LE PLAISIR DANS L'APPRENTISSAGE DES LANGUES Un défi méthodologique*. Perugia.

[39]   B. Review, "Handbook of Research on Educational Communications and Technology," *Educ. Technol.*, vol. 13, pp. 260–263, 2010.

[40]   A. Dhir, "Examining the Educational User Interface , Technology and Pedagogy for Arabic Speaking Children in Kuwait," *J. Univers. Comput. Sci.*, vol. 19, no. 7, pp. 1003–1022, 2013.

[41]   S. Dale, "Gamification: Making work fun, or making fun of work?," *Bus. Inf. Rev.*, vol. 31, no. 2, pp. 82–90, 2014.

[42]   J. Hamari, J. Koivisto, and H. Sarsa, "Does gamification work?–a literature review of empirical studies on gamification," in *System Sciences (HICSS), 2014 47th Hawaii International Conference on*, 2014, pp. 3025–3034.

[43]   S. Vandercruysse, M. Vandewaetere, F. Cornillie, and G. Clarebout, "Competition and students' perceptions in a game-based language learning environment," *Educ. Technol. Res. Dev.*, vol. 61, no. 6, pp. 927–950, 2013.

[44]   G. G. Smith, M. Li, J. Drobisz, H. R. Park, D. Kim, and S. D. Smith, "Play games or study? Computer games in eBooks to learn English vocabulary," *Comput. Educ.*, vol. 69, pp. 274–286, 2013.

[45]   H.-Y. Sung and G.-J. Hwang, "A Collaborative Game-based Learning Approach to Improving Students' Learning Performance in Science Courses," *Comput. Educ.*, vol. 63, pp. 43–51, 2013.

[46]   A. Karkar, J. M. Al Ja'am, S. Foufou, and A. Sleptchenko, "An e-Learning Mobile System to Generate Illustrations for Arabic Text," *2016 IEEE Glob. Eng. Educ. Conf. EDUCON*, no. April, pp. 184–191, 2016.