Semantic Boggle: A Game for Vocabulary Acquisition

Irina Toma[1], Cristina-Elena Alexandru[1], Mihai Dascalu[1,2], Philippe Dessus[3], Stefan Trausan-Matu[1,2]

[1] University "Politehnica" of Bucharest, Faculty of Automatic Control and Computers, 313 Splaiul Independenţei, RO-60042 Bucharest, Romania
irina_toma@rocketmail.com, elena.cristina.alexandru@gmail.com, {mihai.dascalu, stefan.trausan}@cs.pub.ro
[2] Academy of Romanian Scientists, Splaiul Independenţei 54, 050094, Bucharest, Romania
[3] Laboratoire des Sciences de l'Éducation, Univ. Grenoble Alpes, F38000 Grenoble, France
philippe.dessus@univ-grenoble-alpes.fr

Abstract. Learning a new language is a difficult endeavor, the main encountered problem being vocabulary acquisition. The learning process can be improved through visual representations of coherent contexts, best represented in serious games. The game described in this paper, *Semantic Boggle*, is a serious game that exercises vocabulary. It is based on the traditional word-guessing game, but it brings educational value by identifying semantically-related words. The words are found using the *ReaderBench* framework and are placed in the game grid using a greedy algorithm. The assigned score is computed as the semantic similarity value multiplied by the normalized length of the seed. Our preliminary validation consisted of 20 users, who emphasized its interest and playability.

Keywords: Serious games, Vocabulary acquisition, Semantic games, Word-guessing, Boggle, Language learning.

1 Introduction

When learning a new language, people are overwhelmed by the amount of information they should assimilate. This lowers the learners' confidence in their progress and their ability to communicate in a certain language. In order to overcome this, different techniques for vocabulary acquisition have been developed. Traditional techniques focus on memorizing words, using them for translating sentences and expressing grammatical rules, while newer ones propose visual support (i.e., flashcards), group listening activities, or pre-reading activities [1].

Learners, whatever their learning habits and skills are, memorize words used to express their ideas and understand conversations, develop strategies for coping with unknown words and take responsibility for vocabulary expansion [2]. However, each learning technique should be personalized for each individual, and previous knowledge should be integrated with new concepts through comparison, combination, match-making and visual concept representation [3].

2 Vocabulary Games – Serious Games focused on Language

Vocabulary games help learners assimilate words. This process is influenced by the number of times learners are exposed to a word and to its different definitions. If words are not exercised, they do not end up in long-term memory and are easily forgotten [4]. We believe these two challenges can be solved with serious games that behave as a tool for revising vocabulary, in order to adapt its difficulty level to the learner's background.

A game of particular interest for this paper is *MagicWord* [5], a word searching, Boggle-like game that exercises the inflection forms of one language. An alternative gameplay was probed within the E-LOCAL project (https://e-localcourses.unibo.it/) [6], which focuses on the meaning of words, instead of morphology. The experiments with the latter approach showed the need to personalize the game; therefore, an authoring tool was created, allowing teachers and learners to create their own games with user generated content.

3 Semantic Boggle

Boggle is a word game in which players connect neighboring letters in a 4x4 grid. The game is won by the player who finds the longest words. This is a board game, played for fun, but it can be added educational value by looking for specific words or relations between words. The letter arrangement in the 4x4 grid must be predefined and, at the same time, various versions of the grid must be generated. This is a difficult task to be done manually and *Semantic Boggle*, our serious game, was created for that purpose. The main idea of the game is that the grid is populated with semantically-related words, starting from a given word. The algorithm behind the game is described below.

1. *Choose the starting word* (*seed*). The word is chosen randomly from a given lexicon, with the condition that its word length is less than 5 characters.

2. *Populate the game grid with the seed* (see Figure 1). First, a random cell is selected as the starting point for generating the word. Then, each move is randomly generated, with the conditions that the cells are neighbors and an already filled-in cell cannot be occupied. If the algorithm reaches a dead end, where at least one cell cannot be filled-in, the grid is rolled-back to the previous step.

3. *Find and filter semantically similar concepts*. For this step, we use a dedicated web service offered by *ReaderBench*, a framework designed for advanced text analysis [7, 8], to find semantically similar concepts. This service takes as input: a) the seed (the start word), b) the used language (English or French), c) the employed method (e.g., semantic distances in lexical databases – WordNet and WOLF (Wordnet Libre du Français) –, Latent Semantic Analysis – LSA and Latent Dirichlet Allocation – LDA) and d) the corresponding corpus (e.g., TASA or "*Le Monde*" – http://lsa.colorado.edu/spaces.html). The web service outputs a list of similar words and their semantic similarity score using LSA as a semantic model. For the next step only words less than 8 characters, with a similarity score higher than .3 are retained.

4. *Populate the game grid.* The same greedy algorithm used during the second step is used for the other words. If, by any chance, the word cannot be filled-in in 20 moves, the current word is abandoned. In a next iteration of our game, an improvement will be made, similar to the *MagicWord* grid population algorithm [6]. It consists of identifying a common substring, regardless of its position, in the words from the grid and the current word, enabling the re-usage of letters.

5. *Score words selected by the user*. Each selected word is scored as in equation (1), which takes into account the normalized length of the identified words, as well as their semantic similarity. Words that are not similar to the game seed are scored 0, thus guiding the user to converge towards the initial seed word.

$$Score = (1 + ln(length\_lemma)) \times semantic\_similarity(seed, selected\_word) \quad (1)$$

6. *Finish condition***.** The goal of our game consists of identifying the lexical field of the seed word. Once users discover this seed, they have the option of going to the next round, or continuing the game to find additional semantically-related words.

Graphical User Interface. *Semantic Boggle* is a minimalistic web application, where the user forms words by selecting letters from the 4 x 4 grid. The current word is sent to the *ReaderBench* web server together with the seed in order to compute a semantic similarity score. This latter value is then added to the initial score. For keeping track of the used words, checked words are available in the list under *Used Words* and are highlighted based on the similarity with the seed (see Figure 2).

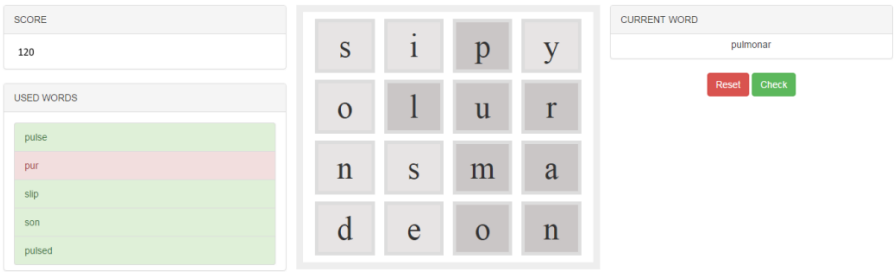

Fig. 2. *Semantic Boggle* – Gameplay print-screen.

4 Results and Discussions

A group of 20 users (60% males) aged 20-27 were asked to play *Semantic Boggle* and provide feedback regarding the gameplay. The users were asked to answer a 10 questions survey with ratings on a 5-point Likert Scale (1 – completely disagree; 5 – completely agree) covering users' perspective on *Semantic Boggle*. Based on all participants' ratings, average intra-class correlation (ICC) was .908 and Cronbach's alpha was .911 denoting a high agreement among the raters. Of particular interest were two questions that demonstrated the adequacy and appeal of our approach: '*Did you enjoy the game?*' (mean = 4.66; SD = 0.478) and '*Did the list of wrong/correct words*

helped you?' (mean = 4.26; *SD* = 0.593). From the three free-answer questions, the general feedback was players appreciated the simplicity of the game and considered it challenging and educative. Two users even agreed it is addictive. As improvements, 50% of the users suggested enhancing the graphical interface, 20% considered adding more complicated words and making them harder to find, and three users proposed personalizing the game based on the background players' knowledge.

Considering the feedback, the most important aspect to improve is the graphical interface, which will be enhanced in terms of usability and design. Second, the algorithm will search for common stubs in order to better fill-in the grid. This will enable teachers to set a list of predefined words to be used, based on the concepts taught and the similar words inferred using *ReaderBench*.

Summing up, *Semantic Boggle* brings a new flavor to the classic *Boggle* game, where an educational aspect is added to the old gameplay. This consists of populating the 4 x 4 grid with semantically-related words, plus scoring learners based on their ability to identify a cohesive context around the seed word. Suggestions for similar words are provided by the dedicated web service available from *ReaderBench* framework, which also includes the capability to assess semantic relatedness. Preliminary validation proved that the concept of the game was appealing to learners. Based on their feedback, a more attractive graphical interface will be implemented, together with a modification of the grid generation algorithm.

Acknowledgment. This work was partially funded by the FP7 2008-212578 LTfLL project, by the 644187 EC H2020 *Realising an Applied Gaming Eco-system* (RAGE) project and by University Politehnica of Bucharest through the Excellence Research Grants Program UPB–GEX 12/26.09.2016. We want also to thank Mathieu Loiseau for all his helpful insights.

References

1. Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Council of Europe ; Cambridge University Press (2000)
2. Přibilová, L.: Teaching vocabulary to young learners. English Language and Literature. Masaryk University, Brno (2006) 46
3. Thornbury, S.: How to teach vocabulary, Vol. 1. Pearson Education India, Longman Essex (2006)
4. Carter, R., McCarthy, M.: Vocabulary and language teaching. Routledge, NY (2014)
5. Loiseau, M., Zampa, V., Rebourgeon, P.: Magic Word – premier jeu développé dans le cadre du projet Innovalangues. ALSIC, 18(2) (2015)
6. Roccetti, M., Salomoni, P., Loiseau, M., Masperi, M., Zampa, V., Ceccherelli, A., Cervini, C., Valva, A.: On the design of a word game to enhance Italian language learning. In: Int. Conf. on Computing, Networking and Communications (ICNC), pp. 1-5. IEEE (2016)
7. Dascalu, M., Dessus, P., Bianco, M., Trausan-Matu, S., Nardy, A.: Mining texts, learner productions and strategies with ReaderBench. In: Peña-Ayala, A. (ed.) Educational Data Mining: Applications and Trends, pp. 345–377. Springer, Cham, Switzerland (2014)
8. Dascalu, M.: Analyzing discourse and text complexity for learning and collaborating, Studies in Computational Intelligence, Vol. 534. Springer, Cham, Switzerland (2014)