Semantic Taboo – A Serious Game for Vocabulary Acquisition

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

[1] University "Politehnica" of Bucharest
313 Splaiul Independenţei, 060042, Bucharest, Romania
E-mail: irina_toma@rocketmail.com, elena.cristina.alexandru@gmail.com, mihai.dascalu@cs.pub.ro, 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
E-mail: philippe.dessus@univ-grenoble-alpes.fr

Abstract. Vocabulary acquisition is considered one of the most important aspects of a language because it defines and constraints communication. In order to overcome this, students and teachers use various techniques meant to improve the learning process and facilitate word acquisition. Using conceptual networks that establish links to prior knowledge is an approach with proven good results. The game presented in this paper, *Semantic Taboo*, is a serious game that helps users enlarge their conceptual network. The game has two modes, one for teachers and one for students. In the teacher module, users have the option of adding playing cards. The theme of the card (the seed) is taken as input, while the list of taboo words is selected from the list of semantically related words provided by the *ReaderBench* framework. The student model has the same game principle as the traditional game, but the list of taboo words and the description of the seed help the learner acquire new words. The preliminary validation consisted of 15 users who were impressed by the game concept and its educational value.

Keywords: serious games, vocabulary acquisition, word-guessing, Taboo game, *ReaderBench* framework.

1. Introduction

The term of serious games is being used to refer to games that have a positive impact on users' skills. This definition includes not only computer games or games that focus on learning, but also those used for medical

treatments, marketing tools, board games, etc. Since serious games reflect such a broad concept, this paper focuses on vocabulary games which help users learn and exercise words in a foreign language.

Learning a new language (L2) is a difficult task for both the learner and his/her teacher because of the large amount of words to be assimilated. Learning strategies are usually divided into two categories (Wu, 2015): incidental learning and intentional learning. *Intentional learning* consists of memorizing words and repeating them until they are familiar. To make this task more interactive, teachers create word lists or cards and link words with their translation in a known language (L1) (Wu, 2014). Proficient learners even apply different learning techniques to create links between concepts and previous knowledge, such as inferring word meaning from the given context. *Incidental learning* consists of performing tasks that are not strictly focused on learning, but have learning as a side-effect. This can be best exemplified by communicating and reading (Laufer & Hulstijn, 2001).

No matter the used method, learners assimilate vocabulary as discrete entities (individual words) or as lexical phrases that are associated pragmatically with other known words. Considering this, a person's lexicon can be assessed as the number of words learned and the connections between the words (Ellis, 2015). The goal of the game described in this paper is to improve the connections between words from the user's vocabulary.

The paper continues with a state of the art on vocabulary acquisition techniques, followed in the third section by a description of conceptual networks. Section four introduces several examples of vocabulary games that follow the previously presented concepts. The next two sections are reserved for the *ReaderBench* framework, as well as the *Semantic Taboo* game description and flow. Results, discussions and conclusions are presented in the last two sections.

2. Vocabulary acquisition techniques

Vocabulary is a fundamental element of all languages, without which learners would be inherently limited when expressing themselves. Having this in mind, the difficulty when learning a new language is acquiring a sufficient number of vocabulary words to become fluent in that language (Kim, 2008). Below we present the most frequently used techniques by

students and teachers to make the acquisition process easier and more rewarding.

2.1 Extensive reading

Reading is part of incidental learning through which users not only enhance their vocabulary, but also improve their spelling and overall language skills. Extensive reading is a pleasurable reading action in which students are encouraged to choose reading materials from the ones recommended by their teacher. The amount of words learned varies considerably based on the frequency of the target words in the text. Moreover, students are more likely to remember a word if they encounter it multiple times (Pigada & Schmitt, 2006).

Another influencer for extensive reading is the target words. These can be partially known, and in this case users understand better how the words are used in different contexts, or unknown. In the latter case, learners either guess the word from the context, or search its meaning in the dictionary. Gu (2003) argued that the dictionary method has better results and that students learn more words and remember them for a longer period of time.

2.2 Extensive Listening

Extensive listening focuses on early learners and consists on teachers reading different materials. For a better understanding and vocabulary acquisition, teachers can explain word meanings as they pass though the text (Brown, Waring, & Donkaewbua, 2008). By combining this technique with extensive reading, students read-while-listen, meaning that the teacher reads the text and the students follow the same text. The reading material can include images, to help the user picture the context. This brings advantages compared to only using one of the previously described techniques. One advantage is that learners can pay attention to the rhythm of the language and understand the segmentation and the coherence of sentences. When reading alone, students can miss the meaning of group of words and try to understand them as separate entities. While the teacher reads, learners comprehend the meaning of the words from their context.

2.3 Word-Focused Tasks

Word-focused tasks promote vocabulary learning through exercises like information gap, sentence writing and word definition (Laufer, 2001). These exercises first expose the user to the target words in different reading materials and afterwards exercise the words. This technique offers a greater advantage compared to extensive reading or listening because students can use the words productively in a conversation or in written texts (Peters, Hulstijn, Sercu, & Lutjeharms, 2009).

Other word-focused activities are list of words and flashcards (Uriarte, 2013). Lists of words are used as a mapping between known words and target words and flashcards have a visual representation, associated pronunciation and spelling. Using these materials, students create a contextual mapping (Wolter, 2006) for the new words that help them integrate the L1 or L2 conceptual network described in detail in the following section. We must emphasize from the beginning the difference between word learning and concept learning: a person can learn words without necessarily understanding them conceptually. Thus, our aim is to support in-depth concept comprehension by defining semantically cohesive contexts. In addition, we must also take into account prior knowledge from student's mind (Bereiter, 2002), therefore greatly impacting the conceptualization of new word acquisitions.

3. Storing vocabulary knowledge in conceptual networks

A conceptual network is a graph with connections between words. Two words are connected if they represent similar concepts (Motter, De Moura, Lai, & Dasgupta, 2002). When learning a new word, usually from a written text, students associate it with existing knowledge, therefore expanding their conceptual network. The learning process is affected by the nature of the word, the text's difficulty, the existing links from the student's contextual network and his/her motivation to learn (Nassaji, 2006). At the beginning of the learning process, students have a nearly empty L2 conceptual network and they use the initial L1 network as a placeholder (Wolter, 2006). This is sufficient in most cases because words usually reflect universal properties. Still, some words require a lexical restructuring of the conceptual network because the concepts from L1 no longer match with those from L2. Wolter

(2006) presents an example of the collocation "furry dog" that is acceptable in English, but "fuzzy dog" or "frizzy dog" are less likely to be encountered, even though the three adjectives are semantically similar.

Contextual networks are defined by breadth and depth. Read (2004) gives the following definitions: breadth is the number of words a person has in his/her vocabulary, while depth represents the quality of the knowledge. Quality does not refer only to the different meanings of the word, but also to pronunciation, spelling, syntactic and semantic relations with other words (Nassaji, 2006).

Both breadth and depth knowledge can be assessed through vocabulary tasks. Since breadth refers strictly to the quantitative aspect of the vocabulary, the tools used are checklists where students mark the word known or unknown, word translations into L1 or synonym identification in a multi-choice test. The evaluation of knowledge depth is more problematic. Read (2000) presents a test called Word-Associate Test (WAT) meant to measure the depth of vocabulary by evaluating the semantic and collocational relations between words.

4. Vocabulary games

From the description of the breadth of evaluation tools, it is understandable that each task can be integrated into one or several games. Games help students learn the target word in a fun and competitive environment and keep them motivated to continue the learning process. On the other side, teachers can revise the vocabulary and grammar in a shorter timeframe and can benefit from potential automated evaluation methods.

Since there is a considerable amount of evaluation tools, vocabulary games have been divided into categories based on the type of actions performed by students (Tuan, 2012):

- *Information gap* – students have to fill-in sentences with missing information;
- *Guessing games* – players have to guess information withheld by other users;
- *Search games* – similar to the information gap games, but the amount of information is larger;
- *Matching games* – students match information with various concepts, pictures or cards.

- *Role-play games* – students play different roles to simulate real life situations and to exercise specific vocabulary.

4.1 Existing games

Vocabulary acquisition is one of the most systematic activities when learning a new language in school contexts; thus, teachers try to make the lessons more interactive and more motivating for the students. This is done by introducing different activities, such as flash cards, word lists or concept maps. Despite this, new concepts should be personalized for each individual student due to different knowledge backgrounds. The next section presents serious games that support users in their vocabulary expansion and are also adaptable to each user's need.

4.1.1 Betty's Brain

Betty (Wagster, Kwong, Segedy, Biswas, & Schwartz, 2008) is a virtual teaching agent whose purpose is to help students research and build models of science phenomena available in the real world. The idea of the game is that Betty answers questions from a virtual mentor and students must analyze her answers and correct her. Betty responds according to a conceptual map and, if questioned, she can provide explanations for her answers which help students identify the texts that were used to generate the responses.

Students are assigned reading materials, which are translated in the conceptual map that Betty uses. This way, the game promotes a shared responsibility as students teach Betty and she answers questions. The conceptual map, that has also a visual representation, organizes the scientific content and helps students understand Betty's reasoning. The game as a whole is motivating for the users because they see their work applied successfully on another entity.

4.1.2 Taboo related games

Taboo is a word-guessing multiplayer game in which one player, the teller, describes the word from a card, while the others try to guess it. The card displays a list of forbidden words that the teller cannot use, thus fostering his or her word knowledge through reformulation and rephrasing; otherwise, the player is disqualified.

This game can be easily adapted and provided in an educational context. Without changing anything in the game, new game cards can be added to exploit to different scientific domains. Capps (2008) and Olimpo, Davis, Lagman, Parekh, and Shields (2010) reflect this purpose-shift (Djaouti, Alvarez, & Jessel, 2011) by using the game for teaching chemistry and biology. This way, students review their knowledge in an entertaining and engaging way, while concurrently promoting a better understanding of the meaning of the terms.

Von Ahn and Dabbish (2004) expanded the idea of the card-game Taboo to a serious game that labels the images available on the internet. This game, called ESP Game, is an online game played by two users. They both see the same image and label what they see. Once the two players write the same label, they receive a new image and the word they agreed on becomes a label for the image. The resemblance with Taboo is that it uses a list of forbidden words that players cannot use as guesses.

Starting from the ESP Game, the same authors proposed a new version of Taboo, called Verbosity (Von Ahn, Kedia, & Blum, 2006). The purpose of the game consisted in constructing a database of "common-sense facts". This seems just a side effect from a player's perspective because the gameplay is centered on word guessing. Verbosity is an online two-player game where one player is the Narrator while the other is the Guesser. The Narrator has a secret word and a list of hints structured as templates. The Narrator fills in the templates with any words, the secret word excepted. The Guesser types any word that (s)he thinks it matches the description and the Narrator tells her/him if it's close to the secret word. The template that the Narrator fills in represents the common-sense fact to be associated with the given word.

5. The *ReaderBench* framework

ReaderBench (Dascalu, 2014; Dascalu, Dessus, Bianco, Trausan-Matu, & Nardy, 2014; Dascalu et al., 2015) is an advanced Natural Language Processing framework that focuses on text cohesion. It was designed as an educational tool to enhance the learning process by assisting both teachers and students. On one hand, teachers can better evaluate the complexity of the textual materials provided to their students, the level of interaction and collaboration between different groups and the students' responses to the

textual tasks (such as summaries and self-explanations). On the other hand, students can identify their level of comprehension and evaluate their learning methods and skills.

The framework exposes a REST API through which its main functionalities can be exploited. From these features, one is of particular interest for our Semantic Taboo serious game: the service that identifies semantically similar words. This service takes as input (Gutu, Dascalu, Trausan-Matu, & Dessus, 2016): a) the target word, b) the used language (i.e., English or French), c) the methods used to calculate the semantic similarity between words (such as semantic distance in lexical databases – WordNet and WOLF (Wordnet Libre du Français), Latent Semantic Analysis, Latent Dirichlet Allocation or word2vec, and d) the corresponding corpus (e.g., TASA or "Le Monde"; http://lsa.colorado.edu/spaces.html). The output of the service is a list of semantically similar words and their similarity score calculated using the corresponding semantic model. The game presented in this paper uses English language as parameter, LSA as the semantic model, and TASA as the corpus.

6. Semantic Taboo

Semantic Taboo is a serious game that follows the same concept as the traditional game, but its purpose is to focus on key words that users must learn or repeat. The target and taboo words are displayed on virtual cards that are either created by teachers or are generated randomly using *ReaderBench*. When the game starts, one user is assigned the host role and receives the game card, while the other players have to guess the taboo word.

6.1 User Interface

Semantic Taboo has an attractive and user-friendly web interface (see Figure 1) developed in JavaScript in which are displayed the four types of interaction the game provides: "*How to play*", "*Create game*", "*Join game*" and "*Add a new card*", as illustrated in Figure 1.

Figure 1. Starting screen of Semantic Taboo

Each game consists of one host and one or more guessers. When selecting "*Create game*", the user is assigned the host role, whereas "*Join game*" assigns the user the guesser role. The host role is to describe the target word without using any of the taboo words and the guessers must find the target word from the description. In order to make this communication easier, the game has a simple chat integrated that is always available to the users.
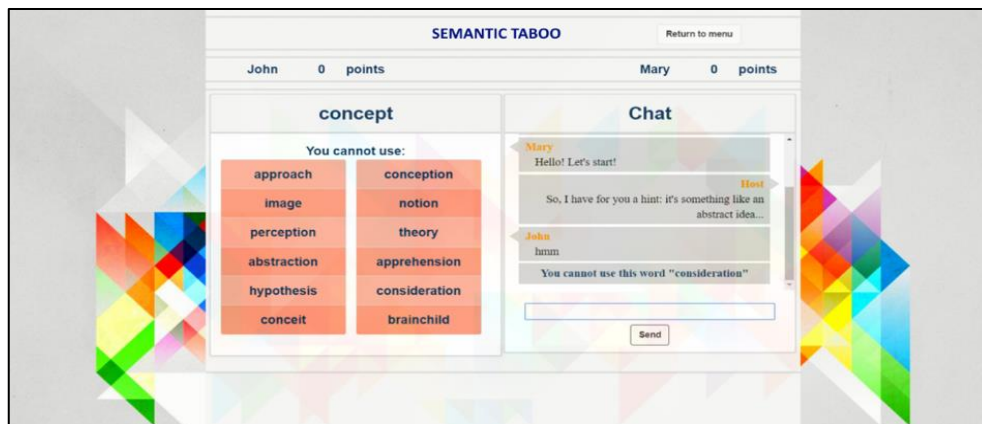

Figure 2. Host interface of Semantic Taboo

The guesser interface is straightforward, consisting of only an input field to send word guesses and the chat window in which users can talk. Each time a guesser introduces an answer in the input field, the word is verified against the target word; if it is a match, the user is awarded a number of points equal to the length of the target word. If the word is incorrect, the

user is penalized by one point in order to discourage the input of random words and to encourage thinking. Once the target word is guessed, all the users are notified in the chat window and the game moves to the next round.

The host interface is more complex, containing the scores in the top part of the screen, the taboo card in the left part and the chat window in the right part (see Figure 2). The host must describe the target word on the card without using any of the taboo words. Each message introduced in the chat by the host is verified against the taboo words.

Another interaction with the game consists of adding a new card (see Figure 3). Cards can be created by teachers or can be generated randomly. Teachers can create cards in a designated game interface that allows them to input one target word and 12 taboo words. This information is stored in the game's database and can be used later in the follow-up games.
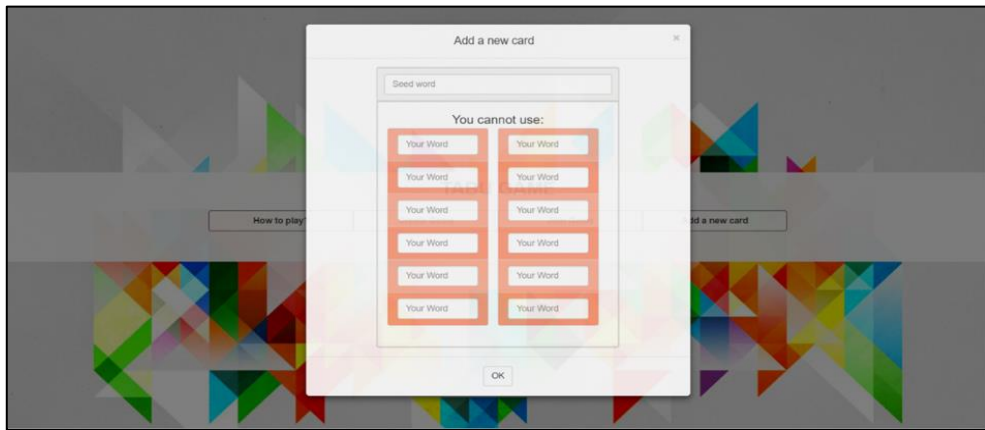

Figure 3. Adding a new card in Semantic Taboo

The second way of creating cards is to automatically generate them using *ReaderBench*. This method uses a dictionary where the entries consist of a lemma and lexemes, the inflectional forms. The taboo word is chosen randomly from the inflectional forms, with the limitation that the length of the word should be more than 7 characters in order to ensure a minimum game complexity. Afterwards, the list of 12 words is populated using the web service provided by *ReaderBench* that generates semantically similar words. Words are filtered based on their similarity score and the first 12 results are chosen as taboo words. If *ReaderBench* returns less than 12 words, the current target word is considered too complex, therefore it is

dropped and the whole generation algorithm is reiterated. The overall flow within the game is depicted in Figure 4.
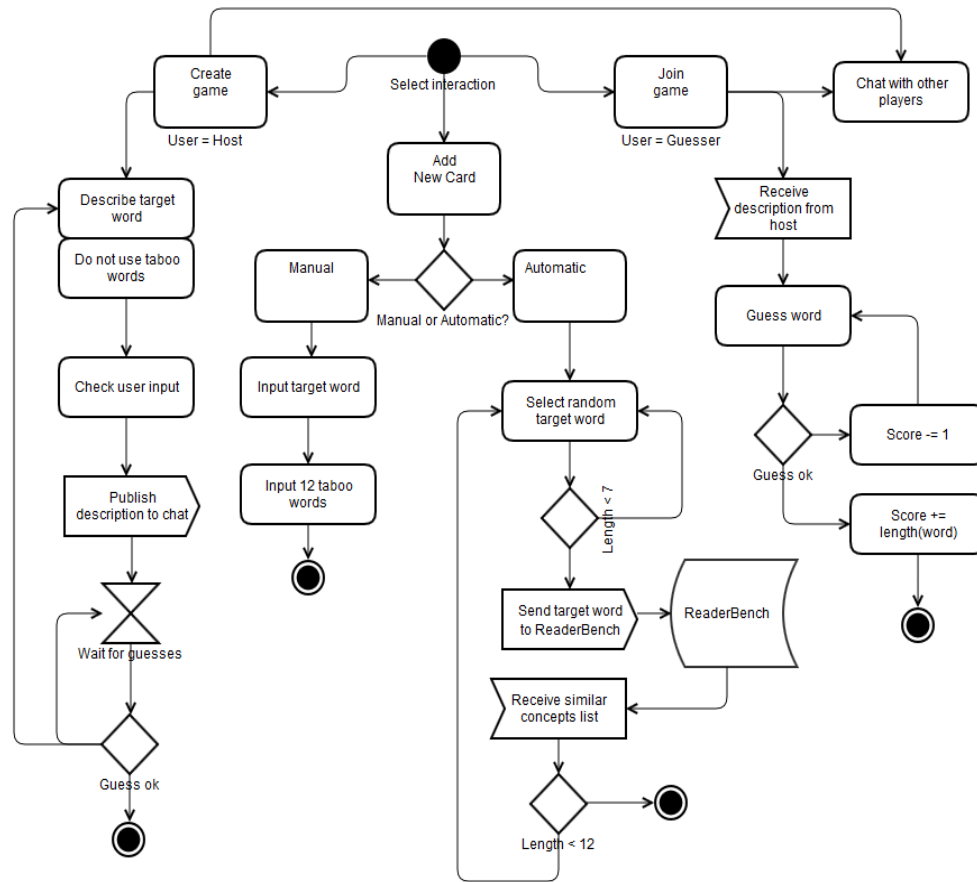
Figure 4. Activity diagram for Semantic Taboo

6.2 Educational scenarios

The educational scenarios in Semantic Taboo are available in both the host and the guesser profiles. The taboo words from the host's interface are similar to the target word, therefore the user can enhance his/her vocabulary with the available words. Also, when describing the target word, the host scans his/her conceptual network to find related concepts to the target word.

On the other side, guessers are presented a definition of a term previously learned, so they investigate their vocabulary to find the target

word. Even though they do not find the target word, all the guesses help
them enhance their conceptual network and make new connections.

7. Results and discussions

A group of 15 students (10 males and 5 females) aged 20-25 were asked to
play the game and answer a 13 questions survey. The survey was divided in
10 questions with ratings on a 5-point Likert Scale (1 – completely disagree;
5 – completely agree) covering the gameplay of Semantic Taboo (presented
in Table 1) and another 3 free answer questions (presented in Table 2).

Based on the participants' ratings, 92% of the users enjoyed playing the
game and 98% considered that the game helped them improve their
vocabulary. Moreover, more than half the users believed that the target
word was easy to find despite the 12 taboo words. As a side effect, users
thought the chat encouraged communication between them, improving their
vocabulary and their writing skills. The *Add new card* functionality was
received with less enthusiasm than we expected as users considered that it
would be easier to suggest words for teachers to be added in taboo list.

Table 1. Semantic Taboo survey

No.	Question	Average Score
1	I enjoyed playing this game.	4.60
2	As Host, it's difficult to explain the target word.	2.13
3	I think that the, „How to play" section is detailed.	4.73
4	I think that this game helped me improve my vocabulary.	4.80
5	This game is very competitive.	4.40
6	The Chat helped me find the word.	4.93
7	I exercised my communication skills.	4.40
8	This game kept me motivated to get a higher score.	4.13
9	The „Add a new card" functionality is useful	3.53
10	The User Interface is nice.	4.80

From the free answer questions, users emphasized the competitive aspect of
the game and seemed satisfied with the chat functionality. They also
enjoyed the colorful interface and the simplicity of the game. On the other
hand, while playing the game with random cards, half the users came across
unknown words and suggested adding definitions for unknown words.
Considering the feedback, the first item to address consists of improving the
Add new card functionality so that words are suggested to teachers. As

teachers input the target word, the application will trigger a call to *ReaderBench* to provide similar concepts. Teachers will have the option to choose words from the similar concept list, or input their own. Also, to double check the taboo words, automated calls will be made to calculate the similarity between the target word and each taboo word. Words with a low similarity score will be highlighted in the interface.

Table 2. Semantic Taboo free answer questions

No.	Question
1	What did you like?
2	What would you improve?
3	What can help you improve your skills while playing this game?

A second improvement is to create an online notebook for each user, host or guesser, where a user can input words definitions and create concept maps to better reflect the current context. Moreover, a dictionary functionality will be added for both for the host and the guessers in order to extend even further the benefits of the online notebook. For the host, each of the taboo words will display a tooltip with the definition of the word and possible usages. The host will have the option to save the definitions to their online notebook. On the other hand, the guesser will be able to see the definition for each inputted guess of the players and for the target word, once this is found.

One last improvement to implement is the option to display the taboo words to the guessers at the end of the game in order to study other similar concepts. This feature will be enabled by default for the randomly generated cards and will be optional for the manually created cards.

8. Conclusion

Summing up, Semantic Taboo follows the footsteps of the initial game, but brings educational value by providing the list of taboo words with semantically similar concepts. This way, students improve their vocabulary and enhance their conceptual network with new terms and relations.

The game also provides a method to create new cards, either by generating them automatically using *ReaderBench* framework, or by inputting them manually. An enhancement for the manual method is to suggest the taboo words and double check the inputted words similarity

coefficient to the target word.

Based on the feedback collected from the users, the game has an enjoyable interface which is both easy to play and competitive. The chat option has a beneficial side effect; even though it was created just for displaying messages, students use it in order to communicate and share opinions, therefore improving their vocabulary and writing skills.

Acknowledgments

References

Bereiter, C. (2002). *Education and mind in the knowledge age*. Mahwah, NJ: Lawrence Erlbaum Associates.

Brown, Ronan, Waring, Rob, & Donkaewbua, Sangrawee. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a foreign language, 20*(2), 136.

Capps, Ken. (2008). Chemistry taboo: An active learning game for the general chemistry classroom. *J. Chem. Educ, 85*(4), 518.

Dascalu, M. (2014). Analyzing discourse and text complexity for learning and collaborating, Studies in Computational Intelligence (Vol. 534). Cham, Switzerland: Springer.

Dascalu, M., Dessus, P., Bianco, M., Trausan-Matu, S., & Nardy, A. (2014). Mining texts, learner productions and strategies with ReaderBench. In A. Peña-Ayala (Ed.), *Educational Data Mining: Applications and Trends* (pp. 345–377). Cham, Switzerland: Springer.

Dascalu, M., Stavarache, L. L., Dessus, P., Trausan-Matu, S., McNamara, D. S., & Bianco, M. (2015). ReaderBench: The Learning Companion. In *17th Int. Conf. on Artificial Intelligence in Education (AIED 2015)* (pp. 915–916). Madrid, Spain: Springer.

Djaouti, Damien, Alvarez, Julian, & Jessel, Jean-Pierre. (2011). Classifying serious games: the G/P/S model. *Handbook of research on improving learning and motivation through educational games: Multidisciplinary approaches, 2*, 118-136.

Ellis, Rod. (2015). Understanding Second Language Acquisition 2nd Edition-Oxford Applied Linguistics: Oxford university press.

Gu, Peter Yongqi. (2003). Vocabulary learning in a second language: Person, task, context and strategies. *TESL-EJ, 7*(2), 1-25.

Gutu, G., Dascalu, M., Trausan-Matu, S., & Dessus, P. (2016). ReaderBench goes Online: A Comprehension-Centered Framework for Educational Purposes. In A. Iftene & J. Vanderdonckt (Eds.), *RoCHI 2016* (pp. 95–102). Iasi, Romania: MATRIX ROM.

Kim, YouJin. (2008). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language learning, 58*(2), 285-325.

Laufer, Batia. (2001). Reading, Word-Focused Activities and Incidental Vocabulary Acquisition in a Second Language. *Prospect, 16*(3), 44-54.

Laufer, Batia, & Hulstijn, Jan. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*(1), 1-26.

Motter, Adilson E, De Moura, Alessandro PS, Lai, Ying-Cheng, & Dasgupta, Partha. (2002). Topology of the conceptual network of language. *Physical Review E, 65*(6), 102.

Nassaji, Hossein. (2006). The relationship between depth of vocabulary knowledge and L2 learners' lexical inferencing strategy use and success. *The Modern Language Journal, 90*(3), 387-401.

Olimpo, Jeffrey T, Davis, Shannon, Lagman, Sarah, Parekh, Raj, & Shields, Patricia. (2010). Learning can be all fun and games: Constructing and utilizing a biology taboo wiktionary to enhance student learning in an introductory biology course. *Journal of Microbiology & Biology Education: JMBE, 11*(2), 164.

Peters, Elke, Hulstijn, Jan H, Sercu, Lies, & Lutjeharms, Madeline. (2009). Learning L2 German vocabulary through reading: The effect of three enhancement techniques compared. *Language learning, 59*(1), 113-151.

Pigada, Maria, & Schmitt, Norbert. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a foreign language, 18*(1), 1.

Read, John. (2000). *Assessing vocabulary* (Vol. 1): Cambridge University Press.

Read, John. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined. *Vocabulary in a second language: Selection, acquisition, and testing, 10*(1), 209-227.

Tuan, Luu Trong. (2012). Vocabulary recollection through games. *Theory and Practice in Language Studies, 2*(2), 257.

Uriarte, Adrian Bello. (2013). Vocabulary Teaching: Focused Tasks for Enhancing Acquisition in EFL Contexts. *Mextesol Journal, 37*(2).

Von Ahn, Luis, & Dabbish, Laura. (2004). Labeling images with a computer game. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 319-326). Vienna, Austria: ACM.

Von Ahn, Luis, Kedia, Mihir, & Blum, Manuel. (2006). Verbosity: a game for collecting common-sense facts. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 75-78): ACM.

Wagster, John, Kwong, Henry, Segedy, James, Biswas, Gautam, & Schwartz, Dan. (2008). Bringing CBLEs into classrooms: Experiences with the Betty's Brain system. In

ICALT'08. (pp. 252-256): IEEE.

Wolter, Brent. (2006). Lexical network structures and L2 vocabulary acquisition: The role of L1 lexical/conceptual knowledge. *Applied Linguistics, 27*(4), 741-747.

Wu, Qun. (2014). A Rote Strategy in Memorizing Vocabulary for ESL Learners. *Procedia-Social and Behavioral Sciences, 143*, 294-301.

Wu, Qun. (2015). Designing a smartphone app to teach English (L2) vocabulary. *Computers & Education, 85*, 170-179..