Formalism for a language agnostic language learning game and productive grid generation

Sylvain Hatier and **Arnaud Bey**
Univ. Grenoble Alpes
Litt&Arts
F-38000 Grenoble
sylvain.hatier@univ-grenoble-alpes.fr
arnaud.bey@univ-grenoble-alpes.fr

Mathieu Loiseau
Univ. Grenoble Alpes
Lidilem
F-38000 Grenoble
mathieu.loiseau@univ-grenoble-alpes.fr

Abstract

In this article, we describe the modifications of MagicWord, a language learning game focused on accuracy, in order to allow the integration of new languages. We first describe the motivations behind the design of the game. Then we explain the modifications performed before exploring the consequences both game-wise and language learning-wise.

In order to improve their replay-value, language learning games need to rely on language resources of diverse complexity depending on their rules and objectives. In this paper, we tackle the issue of providing multi-language resources for a language learning letter game, MagicWord. Before exploring the technical difficulties as well as their intricacies both in terms of language representation, learning and gaming, we will explain the game, its objectives and the design process.

1 Issues of Game-Based Language Learning

Game-based learning gained momentum in the last decade to become a hot topic (Sharples et al., 2013, 29–31) with promises of improved motivation and self-esteem (Cerezo, 2012, 134), and hopes good pedagogy (Gee, 2003; Oblinger, 2004). At the same time that serious games have fostered high hopes, they also brought criticism regarding their actual learning outcomes (Girard et al., 2013) and even their ludicity (Lavigne, 2014; Bruckman, 1999; Söbke et al., 2013).

To us, rather than questioning the concept of serious game, these criticisms underlie the difficulty of creating such games. In this article we will not delve into the complexity and intricate viewpoints on the concepts of “game” and “play” and settle for Brougère’s utterance: “Gaming is a dual reality which interweaves a gaming structure and a playful attitude”[1](Brougère, 2012, 127). The importance of the “playful attitude” inside this sentence underlines one of the central issues towards the creation of a serious game: to be a serious game, the object has to be a game. And to be one, it needs to provoke in the learner a playful attitude. In other words:

> “A good rule of thumb for determining the degree to which a CALL activity is a game [is] the degree to which students want to play it for the pleasure it brings rather than for some external reason. What a teacher or courseware designer calls an activity is not important; it is how the learner views it that will determine whether it is used as one.” (Hubbard, 1991, 221).

Another issue of the design of serious games, that could serve as an explanation of the previous issue, is the cost of developing video games (espe-

[1]“Le jeu est une réalité double qui articule une structure ludique (l’aspect *game* ou *gameplay*) et une attitude ludique (le *play*).” (our translation)

cially those to which the learners are accustomed to playing[2]). Even casual games reportedly require a budget between 100 000 and 1 million dollars (Casual Games Association, 2015), successful games costing more than 500 000 dollars (Handrahan, 2014).

2 Design strategy

In order to try to overcome this issue we have resorted to the following design strategy (Zampa et al., 2017) in various projects:

- adapt Söbke, Bröker and Kornadt's strategy (2013) and select successful commercial-off-the-shelf games which inherently rely on some language competence;
- make sure they allow replayability through generic game mechanics that can be interfaced with language ressources;
- adapt these mechanics so that the language element at the core of the game is made more accessible to the learner (trying not to undermine the playfulness by doing so);
- all this done through various iterations to cut the cost and allow us to get feedback from learners and teachers.

3 Issues of multilingualism in a letter game

MagicWord is one of the (open-source) games designed and implemented using this strategy. Before entering the details of the present iteration of its development, it seems necessary to document the previous stages of the project.

3.1 MagicWord v1

MagicWord is a word game based on the same set of *metaludic* rules (pertaining to a game genre) (Silva Ochoa, 1999, 277) as games like Boggle, Ruzzle, Wordament (and others): 16 letters are set in a 4×4 grid; the goal of the game is to create words (or more precisely word-forms); to do that, the player uses contiguous letters, in every direction (up, down, left, right & diagonally, *cf.* fig. 1), using each letter cell at most once per word form.

We chose this set metaludic rules, because they allow replayability, many forms of the game are successful and those rules rely on the players vocabulary. But what interested us most is that good players tend to try to find all the inflected forms associated with the lemmas they find in the game. And in learning languages with rich stemming mechanisms, such as French, Spanish or German, the learning of the various forms can be considered tedious (Castañeda and Cho, 2016, 1195).

The first version we created was available in Italian, French and English, allowed players to engage in duels. Some games allowed free play and others came with a constraint, that when respected granted extra points (*e.g.* English words ending with "-er"). This version was presented to teachers in a focus group (Montaufier, 2016) and later tested against another version devised in collaboration with the University of Bologna that focused solely on vocabulary (Roccetti et al., 2016). This experiment (Loiseau et al., 2016) and the focus group allowed us to conclude that:

- Our version with duels between two players was better received *playful attitude*-wise than the one player Bologna version;
- That both learners and teachers saw the potential of the game in terms of lexicon rather than in terms of inflections.

3.2 MagicWord v2

Based on this feedback, we created a new version of MagicWord, that built upon the first version but added new rules and functionalities to expand the affordances of the game.

3.2.1 "Massive" games

Without erasing the duel mode, we decided to create a mode where players would compete against the whole community on the same grids in order to create more emulation within the class.

3.2.2 C-c-c-combo

Considering the fact that the *normative* rules (which are followed by experimented players of the game) (Silva Ochoa, 1999, 277) — in this case, trying to find as many forms of each word as possible — are hardly accessible to teachers and learners, we decided to make them part of the *constitutive* rules (explicit rules of a specific game) (Silva Ochoa, 1999, 277) of our game and added a new way to earn points: "combos" (*cf.* fig. 1, left). Combos, short for "combination", are triggered whenever the player selects two words in a

[2] See for instance the notion of AAA game.

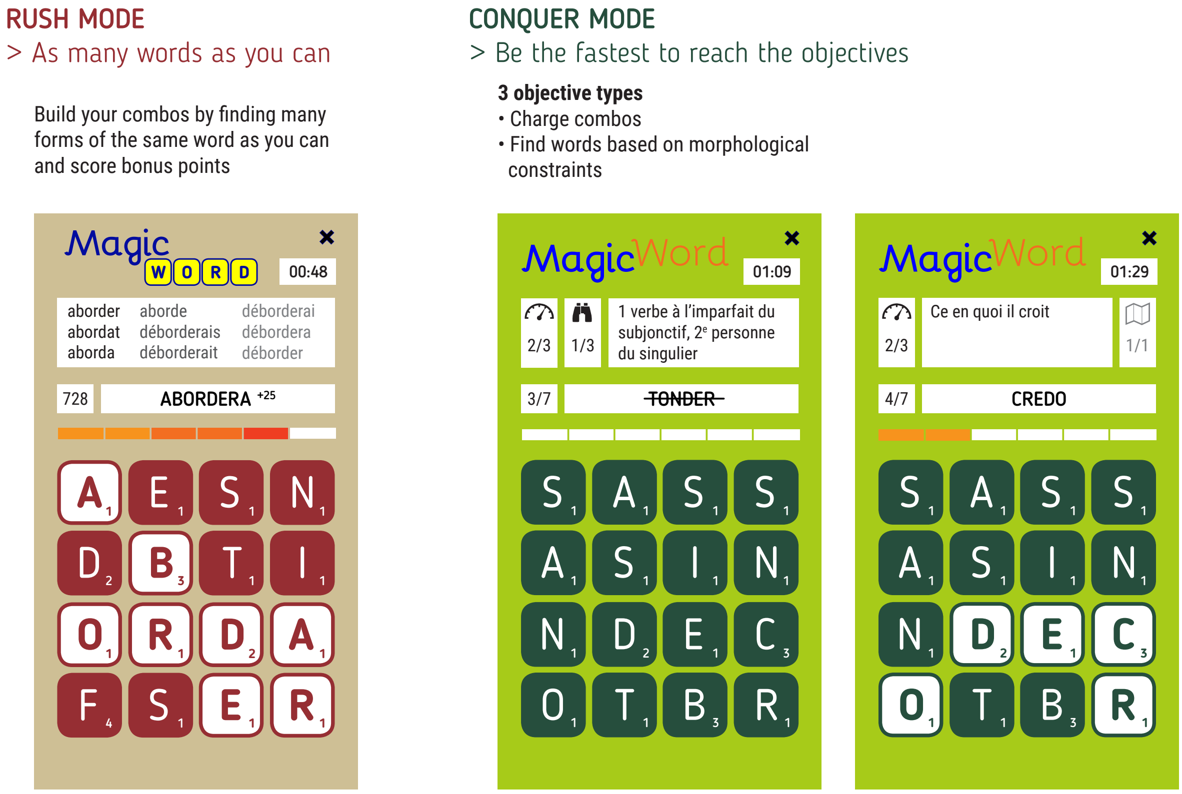

Figure 1: Presentation of the game modes introduced in MagicWord v2 (Loiseau et al., 2017)

row that stem from the same lemma (*e.g.* "play", "playing" in English). The bonus is increased with each new form and broken, whenever the player selects a form that does not exist or that is not linked to the same lemma.

3.2.3 "Conquer" mode and authoring tool

We did not want to overlook the lexical aspect of the game and wanted to give teachers more control over the grid when they wanted to work on specific lexicon items. So we used the algorithm created for the Bologna version (Roccetti et al., 2016) to allow a mode in which players would not score points based on how many words they would find in a limited time, but how little time they would need to find specific words (*cf.* fig. 1, right). Three types of objectives are available in this mode. The user can be asked to: find a word based on a clue written by another user; realize a combo with a minimum of `n` forms; or find `n` forms based on their morphological category. This type of play comes with an authoring tool that allow users to create their own "conquer" grids (or sets of grids).

3.3 Towards a generic multilingual game

While the first version only used a list of inflected forms, the functionalities of version 2 require links between inflected form and their lemmas (*cf.* Combo, section 3.2.2) and access to the morphological features of each inflected form (both machine interpretable and human readable, for the authoring tool). The development time alloted to the project forced us to base the lexicon on the traits of the French language, thus making the structure hardly usable for other languages.

Based on this assumption we undertook the task of creating a data structure and the associated software that will allow administrators to import their own lexicon into MagicWord. The issues raised by this task are manyfolds and intertwine the linguistic nature of the material handled by MagicWord, its game nature, the learning objectives and the overall usability of the software.

4 Updating the datastructure

The first issue is to provide a lexicon structure that will make the game as open as possible, widely and easily used.

4.1 Formalism for the lexicons

We therefore resorted to a rather standard formalism for the lexicon, to wit, tab-separated values (TSV). Given that the lexicon files are encoded in UTF-8, the game should be able to handle most alphabets.

A lexicon row is organized in three columns :

- **Form**: the first column represents the lexical form.
- **Root**: the second column contains a label which connects forms among themselves and which is used for the combo rule (*cf.* section 3.2.2) that was also revamped to make it more generic. We named the column "root" but it can used for anything and the displayed name can be customized, depending on the pedagogical intents, and the language involved. To make full use of the combo rule, that columns could contain the lemma for inflected languages, but it could also, for example, contain an archilexeme, a root word, or even the phonetic transcription if the aim is to work on homophony.
- **Features**: The third column contains features (morphological, grammatical or other), organized in label-value pairs, separated with semi-colons. There are no constraints about the content, the software takes the input as it is formulated.

As a consequence, some perl scripts have been developed to format lexicons from different sources[3] into our formalism. This makes the integration of new languages easier. It is even possible to have several lexicons for a same language coexisting in a MagicWord instance. An administrator can thus provide teachers and learners with a lexicon putting the emphasis on basic vocabulary for instance and another one focusing verbs conjugations, etc.

The "features" column is associated with their transcription in natural language. This is necessary to provide the learner with information on the forms found, but also for anyone who would create a grid for a "conquer" game (*cf.* section 3.2.3) and add a morphological objective. It is planned to add a nesting dimension to the features, which would allow authors and players to adopt a finer approach, manipulating feature classes and subclasses. We also intend to make features label and value translatable in every interface language, either in the specifications file (*cf.* section 4.2) or in the administration panel.

4.2 Game/lexicon specifications

Having defined a generic formalism for the import of the lexicon does not make it useable for the system. Indeed, the game relies on various linguistic elements that will constitute metadata for the import of the lexicon. We will describe in this section the additional information necessary for the system to successfully create a game based on the imported lexicon. This information is to be provided to the system in a text file.

4.2.1 Character rewriting rules

Considering the way words are constructed in the game (*cf.* section 3.1), if the alphabet used contains to many different signs there is a high probability that the number of words to find in grids will be lower, thus making the game more difficult to play. For instance, in French, most letter games (crossword puzzles, boggle, scrabble, etc.) traditionally ignore diacritics. In these games "E" represents at the same time all variations of "E" (case & diacritics): "e", "E", "É", "é", "È", etc.

To allow such behavior, the system can be provided with a set of rules to rewrite the forms only using a subset of the alphabet of the lexicon. After rewriting occurs, whenever the game is played all characters are displayed in uppercase in the interface.

The system gives the administrator the possibility to declare a set of rules that to any given string of letters (made of 1 to n letters) associates a replacement string (empty, single letter or n-gram). E.g. a simple rule to rewrite uppercase 'Œ' and "œ" to the lowercase "oe" would be `RW:Œ,œ=oe`.

Administrators must be aware that the rules are applied in the same order as they are written in the file to avoid side effects. But that gives them the power to shift the balance between accuracy and productivity (number of words in the grids) based on their learning objectives. They can even create 1337 grids if they carefully craft the rewriting

[3] Use of free lexical resources:

- English lexicon: dela (Courtois, 2004);
- French lexicon: morphalou 3 (ATILF, 2019);
- Russian, Spanish, Galician, French & English lexicons: FreeLing (Padró and Stanilovsky, 2012).

rules[4], thanks to the character rewriting rules.

4.2.2 "Rush" mode scoring

In the same way that scrabble is scored differently from one language to the other, MagicWord needs to adapt its scoring system based on the loaded lexicon. The scoring system works around three components to associate to each inflected form found a certain number of points added to the user's score:

- **Letters**: each letter has an inherent value in the game, the base of the word score is the sum of the letters' values;
- **Wordlength**: longer words score more points than shorter words;
- **Combo**: a bonus is attributed if the word is part of a combo and in which position (the 5^{th} word in a combo scores more points than the 3^{rd}).

All these number of scores need to take into account the language. For instance, the English language leaves very little room for (inflection related) combos compared to French, Spanish or German and this should be addressed by putting more weight on the first words in a combo. The frequency of the letters can also be taken into account in the same way as in Scrabble. The length can be neutralized or not depending on the decisions of the administrator, who can specify all the values.

Samples of rules about scoring are given in the specification file model within the game. The administrator can refer to this file in order to write his own rules.

The default value for each **letter** is one point if not defined in the specification file. The interface also allows the administrator to change those values in the administration lexicon section.

Further work is currently planned to attribute automatically a number of points to each letter depending on its frequency in the 'form column' of the lexicon file (after the application of the rewriting rules). The most frequent letters should be associated to the lowest score and vice versa.

By default, the "wordlength" score is set at the size of the word minus one (*i.e.* a four letter word has a "wordlength" score of three). It is also possible to parameter those values in the lexicon specification file but also in the application in the lexicon administration section.

Likewise combo bonus values — established for any length from 2 to 7 words — can be modified by the administrator either through the lexicon specification file or online by using the lexicon administration interface.

By defining a generic formalism both for the entries of the lexicon and the metadata that will allow to create a game out of it, we have explained the influence and control over the game that is granted to the administrator through the import of a new lexicon. Still, the playability of the game depends on the ability of the system to provide grids with sufficient forms available.

4.3 Grid generation

One of the attractive game features of Ruzzle and Wordament is that they are "*fast-paced*"[5][6]. In order not to ruin one of the central aspects of such games, the grid generation should therefore produce letters configuration ensuring a minimum of foundable forms thus keeping grid interest and playability. It is especially important for learners not to feel discouraged. Another issue is to find strategies that will not limit exaggeratedly the coverage of the lexicon.

In this section, we present our strategy to generate grids with sufficient forms available. Our strategy is based on bigrams. In order to quantify our algorithm, we introduce two metrics computed across multiple grids:

- **productivity**: average number of forms and/or combos available in a grid;
- **diversity**: number of distinct lemmas available across generated grids.

4.3.1 Use of bigrams

The letters configuration should depend on the language and *i.e.* the lexicon used. We thus resorted to the use of bigrams. The rationale behind this choice is that depending on their position in the grids each letter is part of 3 (corner), 5 (edge) or 8 (center) bigrams. Is it complicated to have full control over the content of the grid, yet if the bigrams represented in the grid are frequent in the

[4] That will not include the words invented by the "*leet*" community but just the way words are transcribed

[5] "Ruzzle is a fast-paced and addictively fun word game" (Presentation of the game by the mag interactive).

[6] "Wordament is rather fast-paced" (Game review).

language, there is higher probability that the grid contains more words.

In consequence, throughout lexicon import, we list every bigram that occur at least once in a form and calculate bigrams frequency. Then, when a grid is generated, the algorithm organizes the letters in the grid, after randomly drawing them. Rather than resorting to the “scrabble bag” algorithm, like in v1 and v2, we decided to weigh the drawing of letters based on the frequency of bigrams. Once the first letters have been disposed in the grid, letters that are part of more frequent bigrams involving existing letters in the grid are more likely to be selected.

One of the issues of such an algorithm is that the sheer diversity of existing bigrams might produce noise regarding our objectives (high productivity without lowering too much diversity). We therefore introduce a frequency threshold[7] under which bigrams are ignored[8]. For example, if the threshold is set to 5, the 5% less frequent bigrams will disregarded. This has consequences on the grids:

- a positive side effect is the exclusion of parasitic characters (mostly due to encoding issues);
- less frequent bigrams withdrawal (thus augmenting the probability that the bigrams in the grid are used in more forms);
- frequent bigrams occurring in frequent morphemes of the specific lexicon are more represented in grids. As a consequence, inflection bigrams[9] are also more represented (which is one of our objectives for languages with rich inflection mechanisms, *cf.* section 3.1).

4.3.2 Forms diversity & productivity in grids

We have tested this algorithm in relation with various threshold values with the *Morphalou3* french lexicon (668 993 entries). We chose to make the bigram threshold vary between 0 and 100, selecting 13 values (*cf.* fig. 2). Foreach value, we generate 2000 grids and evaluate productivity (average score per grid) and diversity (overall score for 2000 grids).

[7]The administration panel enables the administrator to temper with this threshold value.

[8]Even if they are ignored, they can still occur through the layout of the grid considering that they share letters with more frequent bigrams.

[9]found in morphemes expressing tense, number, gender, mood, etc.

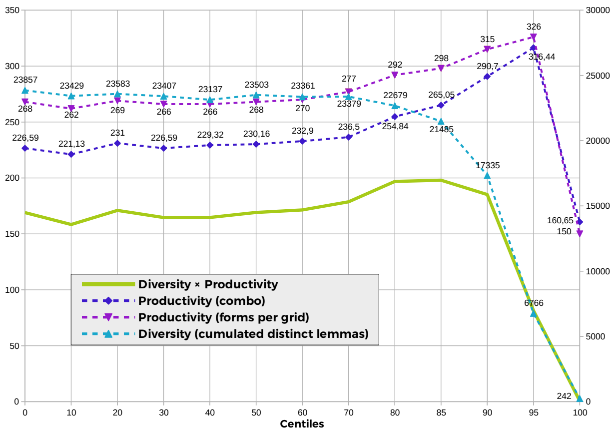

Figure 2: Forms diversity & productivity depending on bigrams centiles

The first element one can note is that combo forms are highly represented. At the ultimate threshold value (99), combos even over come total forms in the grid, which is explainable by the fact that ambiguity can provoke forms to be part of more than one possible combo. For instance, “`lit`” can be found in a combo about the noun “`lit`” (bed: “`lit`”, “`lits`”) and the verb “`lire`” (read: “`lit`”, “`lis`”). In that example, 3 different forms result in 4 possible (and mutually exclusive) combo forms.

As expected, the gap between combo and form productivity reduces as the threshold value increases, which is consistent with the global reduction of diversity as the threshold value increases.

The best diversity over productivity rates are reached between the exclusion of the 80% and 85% less represented bigrams. Over this specific data, this zone is sensitive as the 70% mark initiates a deep dive in diversity that results in five letters (`j`, `k`, `w`, `x`, `y`) being excluded from the game altogether.

With our objectives (working specifically on inflection mechanism), the v3 of MagicWord will be deployed using an 81% threshold value on the morphalou lexicon for “rush” grids generation, in order to maximize productivity, without going to deep in the loss of diversity. It is worth noting that some diversity can be achieved by creating “conquer” grids that will contain specific words.

5 Consequences

From the learner’s standpoint, changes in MagicWord v3 might not seem overwhelming (though the interfaces will be sleeker, especially on mobile devices). All the same we actually underwent a complete overhaul of the data structures that, provided close collaboration between teachers and engineers, can open many doors pedagogically speaking.

The obvious improvement is that MagicWord can now virtually be configured for any alphabetical language. But the generic structure provided to integrate other languages is augmented with modalities that improve drastically the control over the system.

First, all scoring mechanisms can be tuned to focus on certain aspects by putting more or less weight on letters, word length or combos.

More importantly, the data structure, now allows the use of language resources that focus on more diverse phenomena. For instance, if one wants to introduce a more semantic dimension to the game, the resources used to create "Semantic Boggle" (Toma et al., 2017) could be integrated to create semantic combos, thus keeping the fast pace component that some (but not all) players like in Boggle and other derivatives. One can imagine, the same instance of MagicWord could even embark multiple lexicons for the same language to mix rules — *i.e.* viewpoint on the language — inside the same game.

5.1 Future works

In the long run, further improvements will be made to improve administrator control over the game. Two columns might be added to the lexicon.

The first will be scoring information (to be interpreted with a formula provided in the description file). In the long run, frequency lists could be used to refine the scoring process and add a "usage" dimension to it. This will allow a per-word scoring, that could be corpus based (lesser used words scoring more than more widely used words) or even game based (the words more often present in grids and more often found being worth less points).

It should be noted that in the previous iterations of MagicWord, lemmas definitions were automatically retrieved from the wiktionary. In keeping with the genericity, we plan to let administrators define where and how these informations should be collected, by providing urls and regular expressions (or xpath queries). Some institution might have *offline* resources with definitions of the terms in the lexicon written for learners. This information could be provided in the last added column. But the ideal "combo" would be to interface MagicWord with a system of personal lexicon (Mangeot et al., 2016) that would allow to make links between the in-game wordbox[10] with out-of-game more formal activities.

Acknowledgments

The present work has been funded by "Démarre SHS!".

None of this work would have been possible without the many contributiors, most of them volunteers.

References

ATILF. 2019. Morphalou. ORTOLANG (Open Resources and TOols for LANGuage) – `www.ortolang.fr`.

Gilles Brougère. 2012. Le jeu peut-il être sérieux? Revisiter Jouer/Apprendre en temps de serious game. *Australian Journal of French Studies*, 49(2):117–129.

Amy Bruckman. 1999. Can educational be fun? In *Game Developers Conference*, volume 99, page 75–79.

Daniel A. Castañeda and Moon-Heum Cho. 2016. Use of a game-like application on a mobile device to improve accuracy in conjugating spanish verbs. *Computer Assisted Language Learning*, 29(7):1195–1204.

Casual Games Association. 2015. How much does it cost to develop a casual game? FAQ, Casual Games Association.

Catherine Cerezo. 2012. Un serious game junior, vecteur d'estime de soi et d'apprentissages pour des élèves de cm2. *Adolescence*, 30(79):133–143.

Blandine Courtois. 2004. *Dictionnaires électroniques DELAF anglais et français*, volume 24 of *Lingvisticæ Investigationes Supplementa*, page 113–123. John Benjamins, Amsterdam.

James Paul Gee. 2003. What video games have to teach us about learning and literacy. *Computers in Entertainment*, 1(1):20–20.

Coralie Girard, Jean Écalle, and Annie Magnan. 2013. Serious games as new educational tools: how effective are they? a meta-analysis of recent studies. *Journal of Computer Assisted Learning*, 29(3):207–219.

[10]Players can store words of their interest in a personal lexicon called "wordbox".

Matthew Handrahan. 2014. Spil: Casual games “just don’t pay the bills any more”. *gameindustry.biz*.

Philip Hubbard. 1991. Evaluating computer games for language learning. *Simulation & Gaming*, 22(2):220–223.

Michel Lavigne. 2014. Les faiblesses ludiques et pédagogiques des serious games. In *Actes de TiceMed 9*. SFSIC.

Mathieu Loiseau, Cristiana Cervini, Andrea Ceccherelli, Monica Masperi, Paola Salomoni, Marco Roccetti, Antonella Valva, and Francesca Bianco. 2016. *Exploring learners’ perceptions of the use of digital letter games for language learning: the case of Magic Word*, page 277–283. Research-publishing.net, Dublin, Voillans.

Mathieu Loiseau, Virginie Zampa, Racha Hallal, Pauline Ballot, Yoann Goudin, Nadia Yassine-Diab, and John Kenwright. 2017. Gamer (gaming applications for multilingual educational resources).

Mathieu Mangeot, Valérie Bellynck, Emmanuelle Eggerss, Mathieu Loiseau, and Yoann Goudin. 2016. Exploitation d’une base lexicale dans le cadre de la conception de l’enpa innovalangues. In *Actes de la conférence conjointe JEP-TALN-RECITAL 2016*, volume 9 : ELTAL, page 48–64, Paris. ATALA/AFCP.

Agnès Montaufier. 2016. Magic word : comment intégrer le jeu dans une classe de langues ? Master’s thesis, Grenoble.

Diana Oblinger. 2004. The next generation of educational engagement. *Journal of Interactive Media in Education*, 2004(1):1–16.

Lluís Padró and Evgeny Stanilovsky. 2012. Freeling 3.0: Towards wider multilinguality. In *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC’12)*, page 2473–2479, Istanbul, Turkey. European Language Resources Association (ELRA).

Marco Roccetti, Paola Salomoni, Mathieu Loiseau, Monica Masperi, Virginie Zampa, Andrea Ceccherelli, Cristiana Cervini, and Antonella Valva. 2016. On the design of a word game to enhance italian language learning. In *2016 International Conference on Computing, Networking and Communications (ICNC)*, Kaui. IEEE Communications Society.

Mike Sharples, Patrick McAndrew, Martin Weller, Rebecca Ferguson, Elizabeth FitzGerald, Tony Hirst, and Mark Gaved. 2013. Innovating pedagogy 2013: Exploring new forms of teaching, learning and assessment, to guide educators and policy makers. Technical Report 2, The Open University, United Kingdom.

Haydée Silva Ochoa. 1999. *Poétiques du jeu. La métaphore ludique dans la théorie et la critique littéraires françaises au XX*e *siècle.* Thèse, Université Paris 3 — Sorbonne Nouvelle, Paris.

Heinrich Söbke, Thomas Bröker, and Oliver Korndt. 2013. Using the master copy — adding educational content to commercial video games. In *The Proceedings of The 7th European Conference on Games Based Learning*, volume 2, page 521–530, Reading. Instituto Superior de Engenharia do Porto (ISEP), Academic Conferences and Publishing International Limited.

Irina Toma, Cristina-Elena Alexandru, Mihai Dascalu, Philippe Dessus, and Stefan Trausan-Matu. 2017. Semantic boggle: A game for vocabulary acquisition. In *Data Driven Approaches in Digital Education*, page 606–609, Cham. Springer International Publishing.

Virginie Zampa, Nadia Yassine-Diab, and Mathieu Loiseau. 2017. Des jeux et des mots : stratégies de conception et réalisations. *Recherche et pratiques pédagogiques en langues de spécialité*, 36(2).