Vers un réseau lexico-sémantique de grande taille obtenu par *crowdsourcing*
12 ans du projet JeuxDeMots

Mathieu Lafourcade(1), Nathalie Le Brun(2)

(1) Maître de conférences, HDR, université de Montpellier,
(2) chercheur indépendant, Imaginat, Lunel

Résumé : Le projet JeuxDeMots vise à construire une grande base de connaissances de sens commun (et de spécialité), en français, à l'aide de jeux (GWAPs – Games With A Purpose), d'approches contributives, mais également de mécanismes d'inférences. Une dizaine de jeux ont été conçus dans le cadre du projet, chacun permettant de collecter des informations spécifiques, ou de vérifier la qualité de données acquises *via* un autre jeu. Cet article s'attachera à décrire la nature des données que nous avons collectées et construites, depuis le lancement du projet durant l'été 2007. Nous décrirons en particulier les aspects suivants : la structure du réseau lexical et sémantique JeuxDeMots, certains types de relations (sémantiques, ontologiques, subjectives, rôles sémantiques, associations d'idées, etc.), les questions d'activation et d'inhibition, l'annotation de relations (méta-informations), les raffinements sémantiques (gestion de la polysémie), la création de termes agglomérés permettant la représentation de connaissances plus riches (relations à n-arguments).

Mots-clés : réseau lexico-sémantique, base de connaissances, GWAPs (jeux avec but), relations lexicales, relations sémantiques, gestion de la polysémie, représentation des connaissances, connaissances incertaines et floues.

Abstract : The JeuxDeMots project aims to build a large knowledge base of common sense (and also specialty) knowledge, in French, using games (GWAP – games with a purpose), contributory approaches, and inference mechanisms. In this perspective, about ten games have been designed, and each of them allows either to collect specific information, or to check the quality of the data acquired through another game. In this article, we focus on describing the nature of the data collected and constructed since the launch of the project in the summer of 2007. In particular, we discuss the following aspects : the structure of the lexical and semantic network, some types of relations (semantic, ontological, subjective, semantic roles, associations of ideas), activation and inhibition issues, annotation of relations (meta-information), semantic refinements (management of polysemy), the creation of clusters allowing the representation of richer knowledge (n-argument relations)

Keywords : lexical semantic network, knowledge base, GWAPs, lexical relations, semantic relations, polysemy handling, knowledge representation, fuzzy knowledge.

1 Introduction

La construction de ressources lexicales et sémantiques est un problème crucial en TALN [1]. Lorsque cette construction se fait automatiquement à partir de textes, elle peut être biaisée par la nature du corpus choisi (généraliste ou de spécialité), aussi bien que par les méthodes d'acquisition automatique, qui peuvent favoriser l'émergence de certaines informations au détriment d'autres. L'extraction automatique de relations typées entre termes, à l'aide de patrons syntaxiques, se heurte en particulier à l'ambiguïté lexicale et aboutit souvent à un bruit important, nécessitant une correction [2]. L'acquisition de ressources peut également être effectuée de façon manuelle selon plusieurs méthodes, les ressources pouvant être créées : a) par des experts : c'est alors un processus long et coûteux ; b) par des approches contributives, rémunérées ou non, sollicitant des participants attirés par les choses de la langue ; c) enfin par une stratégie consistant à mettre en œuvre des jeux accessibles sur Internet, où les joueurs interviennent sans forcément connaître la dimension contributive de leur implication. Ces jeux permettent d'acquérir des ressources par consensus populaire entre locuteurs natifs. Il s'agit d'un processus rapide et peu coûteux ; cependant, la question de la qualité de la ressource peut légitimement être posée.

2 Acquisition d'un réseau lexico-sémantique avec des jeux

Les jeux visant à construire des ressources restent rares, même si ce domaine est en expansion [3] . Dans la littérature, il est souvent question de *Human Computation* lorsque l'activité (plus ou moins ludique) consiste à substituer des utilisateurs à la machine, pour des tâches en général difficilement automatisables de façon satisfaisante. Si le nombre d'utilisateurs est important [4], il s'agit alors de *crowdsourcing* (« peuplonomie »), que l'approche soit contributive ou ludo-contributive. Il est à noter que les approches contributives rémunérées à la tâche posent un certain nombre de questions quant à leur efficacité et leur éthique, notamment dans le cadre d'*Amazon Mechanical Turk* [5]. Concernant les données de nature lexicale, sémantique ou de connaissance du monde, les exemples d'approches ludiques restent peu nombreux.

D'une façon générale, les données produites par les jeux ne sont pas, à notre connaissance, librement disponibles. Il est possible de tenter d'analyser les mécanismes des approches proposées, mais il reste difficile, voire impossible, de confronter ces approches avec les données effectivement récoltées ou construites.

1. TALN : traitement automatique des langues naturelles.

2. Voir Fabre, Cécile et Bourigault, Didier (2006) « Extraction de relations sémantiques entre noms et verbes au-delà des liens morphologiques », *Proceedings of TALN' 06*, Leuven, Belgique, p. 121-126.

3. Lafourcade, Mathieu ; Le Brun, Nathalie ; et Joubert, Alain (2015) *Jeux et intelligence collective – résolution de problèmes et acquisition de données sur le Web,* ISTE éditions, coll. « Science cognitive et management des connaissances ».

4. Zesch, Torsten et Gurevych, Iryna (2009) « Wisdom of crowds *versus* wisdom of linguists measuring the semantic relatedness of words », *Natural Language Engineering*, Cambridge University Press, p. 25-59.

5. Voir Fort, Karën ; Adda, Gilles ; et Cohen Bretonnel, Kevin (2011) « Amazon Mechanical Turk : Gold mine or coal mine ? », *Computational Linguistics*, 37 (2), p. 2-9 ; et Sagot, Benoît ; Fort, Karën ; Adda, Gilles ; Mariani, Joseph et Lang, Bernard (2011) « Un turc mécanique pour les ressources linguistiques : critique de la myriadisation du travail parcellisé », *Proceedings of Traitement Automatique des Langues Naturelles (TALN)*, Montpellier, France. Accessible sur `https://hal.inria.fr/inria-00617067`, consulté le 6 août 2018.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

Les réseaux lexicaux, qu'ils soient généraux ou spécialisés, constituent, en particulier depuis l'émergence de WordNet [6], une ressource lexicale considérée comme particulièrement précieuse pour les applications de l'informatique linguistique, dont notamment la désambiguïsation lexicale [7]. Dans la plupart des réseaux lexicaux disponibles, les relations ne sont pas pondérées, c'est-à-dire qu'elles sont simplement énumérées, sans aucune indication quant à leur « force » ou leur fréquence. L'introduction d'une telle pondération permet de faire la distinction entre les relations fortes (c'est-à-dire les termes fortement associés), et celles qui sont plus anecdotiques, ces dernières n'étant pas nécessairement moins importantes dans les processus de TALN. Les algorithmes de propagation en désambiguïsation peuvent tirer profit de telles pondérations. Cependant, la question de la détermination des pondérations reste délicate : elles sont fortement reliées à la fréquence d'usage de la relation, mais pas uniquement. Ainsi, un réseau lexical avec des relations pondérées peut être construit avec plus ou moins de bonheur manuellement et/ou par croisement de données avec des procédures plus ou moins automatisées [8] suivies d'une vérification, manuelle également [9].

De plus, les réseaux lexico-sémantiques contiennent souvent un nombre variable de types de relations, généralement ontologiques (hyperonymes) ou lexicales (synonymes, antonymes). HowNet [10] est un exemple de ce type de ressource lexicale (chinois et anglais) construite manuellement et présentant un nombre important de types de relations. CYC [11] illustre un effort majeur de construction manuelle d'une base de connaissances générales fondées sur la logique, dont la forme même peut être assimilable à un réseau sémantique, mais pas ou peu lexical. Enfin, certains réseaux contiennent également des relations simplement associatives, monolingues ou multilingues. Par exemple, BabelNet [12] est un réseau multilingue de grande taille construit par extraction/compilation automatique des informations de Wikipedia.

---

6. Voir, pour EuroWordNet, Miller, George A. ; Beckwith, Richard ; Fellbaum, Christiane ; Gross, Derek et Miller, Katherine (1990) « Introduction to WordNet : an On-Line Lexical Database », *International Journal of Lexicography*, 3 (4), p. 235-244 ; Fellbaum, Christiane et Miller, George A. (dir.) (1998) *WordNet. An Electronic Lexical Database,* The MIT Press ; et Vossen, Piek (1998) *EuroWordNet : A Multilingual Database with Lexical Semantic Networks*, Kluwer, Dordrecht, The Netherlands.

7. Véronis, Jean, et Ide, Nancy (1990) « Word Sense Disambiguation with Very Large Neural Networks Extracted from Machine Readable Dictionaries », *Proceedings of 13th International Conference on Computational Linguistics (COLING'90)*, Helsinki, vol. 2, p. 389-394.

8. Voir Joubert, Alain et Lafourcade, Mathieu (2012) « A new dynamic approach for lexical networks evaluation", *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, Istanbul, Turkey ; et Lafourcade, Mathieu et Le Brun, Nathalie (2017) « Extracting semantic relations *via* the combination of inferences, schemas and cooccurrences », *Proceedings International Conference on Recent Advances in Natural Language Processing (RANLP 2017)*, Varna, Bulgaria, September 2-8, 2015, p. 417-423.

9. Voir Sagot, Benoît et Fišer, Darja, 2008, « Construction d'un wordnet libre du français à partir de ressources multilingues », *Proceedings of TALN 2008*, Avignon, France. Accessible sur `https://hal.inria.fr/inria-00614707/document`, consulté le 29 juillet 2020 ; Lafourcade, Mathieu ; Joubert, Alain et Le Brun, Nathalie (2017) « Si les souris étaient des reptiles, alors les reptiles pourraient être des mammifères ou Comment détecter les anomalies dans le réseau JDM ? », *24 rd French Conference on Natural Language Processing* (TALN 2017), Orléans, France, 26-30 Juin 2017.

10. Dong, Zhendong et Dong, Qiang (2006) *HowNet and the Computation of Meaning*, WorldScientific Publishing Co., London.

11. Lenat, Douglas B. (1995) « CYC : A large-scale investment in knowledge infrastructure », Communications of the ACM, 38(11), p. 33-38.

12. Navigli, Roberto et Ponzetto, Simone (2010) « BabelNet : Building a very large multilingual semantic network », *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, Uppsala, Sweden, p. 216-225.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

C'est au type de méthode reposant sur des « jeux sérieux finalisés » ou « jeux avec finalité d'acquisition de ressources » ou encore en anglais Games With A Purpose (GWAPs [13]), appliqué à ce type de ressource (les réseaux lexico-sémantiques), que nous nous intéressons dans cet article. Nous illustrerons notre propos en détaillant l'expérience du projet JeuxDeMots [14] (noté JDM par la suite) qui vise à produire un réseau lexico-sémantique. Nous analyserons les données collectées et nous en déduirons quelques difficultés rencontrées par les joueurs, mais également dans quelques aspects de la conception des jeux.

Quelques projets en rapport avec le TALN ont exploité avec plus ou moins de succès la démarche ludo-contributive, mais peu se sont maintenus sous la forme de jeux pérennes. Phrase Detective, qui fait partie de ces derniers, vise à produire des annotations de corpus textuels sur divers phénomènes linguistiques, dont notamment la coréférence [15]. Il est à noter de nombreux efforts vers la gamification [16], ou encore la conception d'interfaces homme-machine intelligentes orientées vers la collecte de données *via* des contributeurs humains [17].

Formellement, un réseau lexico-sémantique est un graphe, où les nœuds sont des objets lexicaux (typiquement des termes), reliés entre eux par des relations lexicales ou sémantiques. Ainsi, une relation est un quadruplet : nœud de départ, nœud d'arrivée, type de la relation, poids. Par exemple : *chat r_has_part=190 moustache* (la relation a un poids de 190 – un poids non précisé est implicitement positif). Il est ainsi possible de représenter de façon unifiée et non limitante toutes sortes d'informations. Par exemple, les parties du discours (*part of speech* ou *POS* en anglais), sont eux-mêmes des nœuds, vers lesquels pointent les termes. Pour représenter que le terme chat est

13. Von Ahn, Luis (2006) « Games with a purpose », *Computer*, 39(6), IEEE Computer Society Press Los Alamitos, CA, USA, p. 92-94.

14. Lafourcade, Mathieu, 2007, « Making People Play for Lexical Acquisition », *Proceedings SNLP 2007, 7th Symposium on Natural Language Processing*, Pattaya, Thaïlande. Accessible sur `https://hal-lirmm.ccsd.cnrs.fr/lirmm-00200883/document`, consulté le 29 juillet 2020.

15. Sur PhraseDetective, voir Chamberlain, Jon Poesio, Massimo et Kruschwitz, Udo (2009) « A demonstration of human computation using the Phrase Detectives annotation game », *HCOMP '09 : Proceedings of the ACM SIGKDD Workshop on Human Computation*, p. 23-24 ; et Chamberlain, Jon ; Poesio, Massimo et Kruschwitz, Udo (2008) « Phrase Detectives : A Web-based Collaborative Annotation Game », *Proceedings of the International Conference on Semantic Systems (I-Semantics'08)*, Graz. Disponible sur `https://dces.essex.ac.uk/staff/udo/papers/phrasedetectives.pdf`, consulté le 3 août 2018. Sur la coréférence, voir Hladka, Barbora ; Mirovsky, Jiri, et Schlesinger, Pavel (2009) « Play the language : play coreference », *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers*, p. 209-212.

16. Kuo, Yenling ; Lee, Jong-Chuan ; Chiang, Kai-Yang ; Wang, Rex ; Shen, Edward ; Chan, Chen-Wei et Yung-jen Hsu, Jane (2009) « Community-based game design : experiments on social games for commonsense data collection », *Proceedings of the ACM SIGKDD Workshop on Human Computation, HCOMP'09*, New York, USA, p. 15-22. Von Ahn, Luis ; Kedia, Mihir et Blum, Manuel (2006) « Verbosity : a game for collecting common-sense facts », *In Proceedings of ACM CHI 2006 Conference on Human Factors in Computing Systems*, volume 1 of Games, p. 75-78. Siorpaes, Katharina et Hepp, Martin (2008) « Games with a Purpose for the Semantic Web », *IEEE Intelligent Systems*, vol. 23, no. 3, p. 50-60. Von Ahn, Luis et Dabbish, Laura (2004) « Labelling Images with a Computer Game », *ACM Conference on Human Factors in Computing Systems (CHI)*, New York, NY, USA, ACM Press, p. 319-326. Lieberman, Henry ; Smith, Dustin et Teeters, Alea (2007) « Common Consensus : A Web-based Game for Collecting Commonsense Goals », *Workshop on Common Sense for Intelligent Interfaces, ACM International Conference on Intelligent User Interfaces (IUI-07)*, Honolulu. Accessible sur `http://web.media.mit.edu/~lieber/Publications/Common-Consensus.pdf`, consulté le 6 août 2018.

17. Chklovski, Timothy et Gil, Yolanda (2005) « Improving the design of intelligent acquisition interfaces for collecting world knowledge from Web contributors », *Proceedings of K-CAP* '05, p. 35-42. Mihalcea, Rada (2003) « Open MindWord Expert : Creating large annotated data collections with Web users help », *Proceedings of the EACL 2003, Workshop on Linguistically Annotated Corpora (LINC 2003)*. Accessible sur `https://doctiktak.com/relation-inference-in-lexical-networks-with-refinements.html`, consulté le 6 août 2018.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

un nom, on aura la relation suivante : *chat r_pos Nom: Le nœud Nom:* n'est pas un terme du lexique, mais un symbole indiquant la partie du discours correspondante.

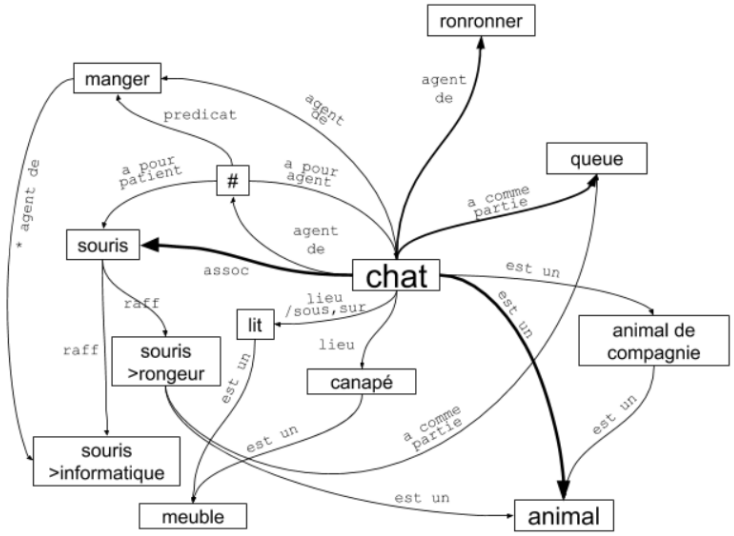


FIGURE 1 – Portion simplifiée du réseau lexico-sémantique JeuxDeMots. Le noeud '#' est un agrégé dont le sens est "verbe: manger ; agent: chat ; patient: souris"

L'intérêt d'une telle structure est que toutes les informations (objets) ou connaissances (rapports entre ces objets) y sont représentables. Les nœuds sont des termes ou des informations diverses (*POS*, lemme, etc.), et ils sont reliés par les relations entre termes (par exemple *lit* est un *meuble*) ou entre un terme et une information (*manger* a comme *POS verbe*).

La notion de graphe est relativement universelle, puisqu'elle existe en mathématiques, en informatique ou encore en sciences sociales. Sans rentrer dans des détails trop techniques, il est possible de concevoir simplement que chercher des informations dans un graphe équivaut à établir un schéma (qui est lui-même un graphe) et à trouver tous les endroits où il peut s'appliquer dans le graphe. Par exemple, si nous cherchons des *animaux blancs*, nous pouvons définir le schéma suivant : *$x r_isa animal & $x r_carac blanc* (un x qui est un animal et qui a pour caractéristique blanc) puis chercher dans le réseau lexico-sémantique tous les nœuds vérifiant ces relations.

3 Quelles propriétés développer pour un jeu lexical ?

Les caractéristiques conditionnant le succès d'un jeu lexical se définissent par les réponses à quelques questions simples : que peut-on demander aux joueurs ? Comment ? Sous quelle forme ? Certains éléments de la grille d'analyse qui suit sont transposables

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

à des jeux non lexicaux. On dispose de peu d'analyses concernant les GWAPs, la plupart restent très générales [18] ou se limitent à décrire ce qui existe [19].

3.1 Pour les joueurs

Pour attirer des joueurs, puis les inciter à revenir participer, un jeu doit être :

— Amusant : présenter un intérêt ludique (n'est-ce pas une évidence ?) au niveau de l'interface afin d'*accrocher* les joueurs, mais plus encore au niveau du contenu afin de les garder est indispensable. L'activité ludique proposée doit éviter la monotonie tout en incluant la répétitivité nécessaire pour que le joueur puisse s'amender. Dans le cas de construction de ressources lexicales et terminologiques, la variété du vocabulaire offre des possibilités de renouvellement quasiment illimitées. La diversité potentielle de parties proposées est encore accrue si la ressource inclut des relations lexico-sémantiques de types variés, des niveaux de langue, des informations en rapport avec les phénomènes de polysémie, etc. L'ambiguïté inhérente au lexique est aussi une source d'intérêt et de surprise pouvant être exploitée avec profit.

— Facile à appréhender : beaucoup de joueurs potentiels ne lisent pas les consignes, il faut donc que le jeu soit intuitif. Dans le cadre d'un jeu lexical, de nombreuses relations sont directement compréhensibles par le locuteur et ne nécessitent généralement pas un bagage linguistique conséquent. Il est à peu près certain que les données dont l'appréhension nécessite une formation linguistique poussée (par exemple, des arbres d'analyse morphosyntaxiques) ne peuvent pas être sollicitées dans ce type de contexte.

— Valorisant : un système de points et de classement entre les joueurs rend en général le jeu plus attirant – la possibilité de compétition entre joueurs incite certains à y consacrer beaucoup de temps, donc à acquérir de l'expérience, ce qui les rend non seulement plus productifs, mais augmente globalement la qualité des contributions. Il est très rare qu'un joueur soit rebuté par la présence de classements, que le cas échéant il peut toujours ignorer. Dans des jeux impliquant une confrontation au système, les joueurs peuvent être mis en situation de se sentir « plus intelligents que la machine » : par exemple, dans Totaki, le jeu de devinettes [20], un grand nombre de parties visent à piéger le système, ou à lui apprendre quelque chose ; les joueurs ont de plus le sentiment (légitime) de jouer un rôle-clé dans le processus d'acquisition des connaissances.

— Évolutif : les joueurs acquièrent de l'expérience en jouant, ils répondent de manière de plus en plus pertinente, à des questions de plus en plus précises. Par exemple, une partie demandant des *idées associées* au terme *viande* est intéressante pour découvrir JDM parce que nombreuses réponses viennent immédiatement à l'esprit ; mais par la suite, les joueurs préfèrent des consignes plus difficiles, comme « Quelles sont les *causes* possibles de *turgescence* ? » ou encore « Quelles sont les *significations* possibles

18. Von Ahn, Luis et Dabbish, Laura (2008) « Designing Games With a Purpose », *Communication of the ACM*, 51(8), p. 58-67.

19. Thaler, S. ; Simperl, E. ; et Siorpaes, K. (2011) « SpotTheLink : A Game for Ontology Alignment ». *In Proc. of 6th Conference on Professional Knowledge Management : From Knowledge to Action*, February 21-23, 2011, Innsbruck, Austria.

20. Lafourcade, Mathieu et Zampa, Virginie, 2009, « PtiClic : a game for vocabulary assessment combining JeuxDeMots and LSA », *Proceedings of CICLing (Conference on Intelligent text processing and Computational Linguistics)*, Mexico.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

du terme *moyen* ? » Il est intéressant que la difficulté augmente avec l'expérience du joueur. Toujours dans un jeu lexical et/ou sémantique, le niveau de difficulté peut être fonction des propriétés lexicales déjà découvertes (et validées) des termes du lexique.

— À parties courtes : plus les parties sont courtes, moins le joueur potentiel hésite à jouer. Les approches lexicales en général, et les réseaux lexico-sémantiques en particulier peuvent être fondés sur des micro-tâches dont la granularité est très fine et le temps de réalisation très faible. Dans le cas de JDM, la durée de jeu prévue par les joueurs (estimée par enquête) pour une session est de l'ordre de 10 minutes, mais la durée effective mesurée avoisine 25 minutes, une partie durant 60 secondes par défaut.

3.2 Pour les concepteurs

Pour les concepteurs d'un jeu dont l'obtention de ressources de qualité constitue le but initial, cette méthode d'acquisition présente un certain nombre de qualités. Les ressources produites sont le fruit de résultats consensuels, partagés par un grand nombre de locuteurs. Le résultat est la constitution d'une base de connaissances générales communes, auxquelles viennent s'ajouter des connaissances plus spécifiques. Le prix de revient des ressources est très faible puisque limité à celui du développement, puis de la maintenance.

Un jeu d'acquisition lexicale doit idéalement s'adresser à un public aussi large que possible. Cependant, la nature même des objets manipulés peut réduire fortement cette espérance. Dans le cas de JDM, 70 % des joueurs ont entre 30 et 60 ans, 60 % sont des femmes, et la quasi-totalité a un niveau d'études supérieur. La promotion du projet JeuxDeMots a été effectuée sur les réseaux sociaux. Des flyers ont aussi été distribués lors de colloques. Il est clair que ce type de diffusion cible un public de niveau d'études post-bac. Par ailleurs, le jeu a également été présenté sur des forums de jeux généralistes, et des retours sporadiques semblent montrer que le jeu fait peu d'adeptes auprès des personnes ayant un niveau d'études inférieur au bac. Les retours montrent aussi que les hommes « accrochent » beaucoup moins au jeu que les femmes, en général. Depuis le lancement du jeu, plus de 6 000 personnes se sont inscrites, sachant qu'il est également possible de jouer sans s'identifier, en « invité ». Quant aux joueurs que l'on peut qualifier de « fidèles », ils sont de deux types : moins de dix joueurs jouent régulièrement, sans interruption, depuis qu'ils ont découvert le jeu (pour certains, depuis l'origine, soit 2008), tandis qu'un nombre nettement plus élevé de personnes ont joué très assidûment pendant une période donnée, accumulant un grand nombre de parties (plus de 200 joueurs ont joué plus de 1000 parties, et plus de 3 800 joueurs entre 100 et 1000 parties, pour un total de plus de 1,54 million de parties jouées, par les inscrits, pour le jeu JeuxDeMots uniquement) pour s'interrompre ensuite. Parmi ces joueurs, certains reviennent sporadiquement, d'autres pas.

On remarquera enfin que, d'un point de vue éthique, solliciter des contributeurs à travers des jeux semble ne pas poser autant de problèmes que dans un cadre rémunéré (Sagot *et al.*, 2011), même si à première vue les tâches semblent très similaires. La question de l'efficacité est également débattue, mais *a priori* la motivation des participants à bien jouer induit la constitution d'une ressource de bonne qualité.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

3.3 La question du renouvellement des joueurs

Le roulement (renouvellement des joueurs) peut être très court, d'autant plus s'il y a phénomène de mode au sein d'un sous-groupe ou d'une petite communauté de joueurs, qui peuvent collectivement abandonner le jeu pour passer à autre chose. Selon notre expérience sur JDM, le roulement est d'environ trois semaines. Nous avons constaté qu'un certain nombre de joueurs reviennent, après une longue période d'inactivité, parfois simplement pour jouer de nouveau, mais le plus souvent avec une attitude différente, plus ciblée, par exemple, pour vérifier avec le jeu TOTAKI (voir 3.2) si le système connaît telle ou telle information.

Un roulement trop rapide des joueurs est un inconvénient : en effet, la majorité des joueurs n'ont pas le temps d'acquérir suffisamment d'expérience dans le jeu pour en exploiter toutes les subtilités, et restent trop souvent à un niveau de jeu sommaire.

Notons que grâce à l'aspect addictif, entretenu par le système de points et de classement qui favorise l'émulation, et à la diversité des thématiques disponibles, une proportion non négligeable de joueurs « fidèles » permet d'acquérir des ressources considérées comme spécialisées parce qu'elles procèdent d'une compétence particulière dans un domaine précis.

4 Quelques jeux du projet JeuxDeMots

La présentation de jeux du projet JeuxDeMots, permet de rendre concrètes certaines des propriétés désirables mentionnées ci-dessus. De plus, l'expérience menée depuis 2008 permet d'évaluer l'impact effectif en termes de quantité et qualité des ressources collectées.

Dans le cadre du projet JDM [21], nous avons développé plusieurs jeux, avec chaque fois des objectifs différents. Nous présentons certains d'entre eux ci-après, en en évoquant les limites. Il est notable que chaque jeu est pourvoyeur d'informations mémorisées dans le réseau lexical selon des modalités bien précises, mais qu'il constitue également un « contre-jeu » en ce qu'il permet de compenser, au moins partiellement, les biais induits par chacune des autres modalités.

4.1 Jeux d'associations

JDM, le tout premier jeu du projet, mis en service en juillet 2007, est entièrement dévolu à la constitution du réseau lexical. La base lexicale a été amorcée avec environ 150 000 termes sans aucune relation entre eux. Au bout de treize ans, au 1er août 2020, le réseau contient plus de 300 millions de relations et 4,1 millions de termes.

21. Décrit en particulier dans Lafourcade, Mathieu, « Making People Play for Lexical Acquisition », art. cité, et dans Lafourcade, Mathieu et Joubert, Alain (2010) « Computing Trees of Named Word Usages from a Crowdsourced Lexical Network », *Investigationes Linguisticae*, vol. XXI, p. 39-56.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

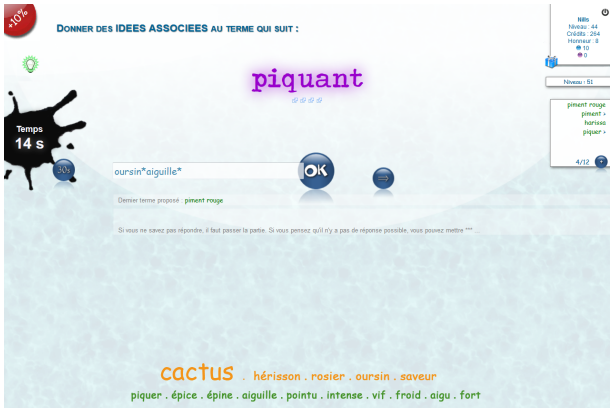


FIGURE 2 – Exemple d'une partie de JeuxDeMots où le joueur doit indiquer à quoi lui fait penser le terme « *piquant* ». Les termes en orange, en bas de l'écran, sont tabous.

JDM est un jeu dit « ouvert », les joueurs devant entrer en un temps limité des propositions à partir d'un terme cible et d'une consigne. Cette modalité d'interaction réduit fortement les *scenarii* d'utilisation, mais est de loin la plus riche de l'ensemble des jeux proposés. Une partie de JDM, sur un terme et un type de relation, se joue de façon asynchrone entre deux joueurs, et de telle sorte que le joueur actif ne connaît pas l'identité du joueur qui a fourni les réponses auxquelles vont être confrontées ses propres propositions, le pseudo de l'adversaire n'apparaissant qu'au moment du résultat. Il n'est donc pas possible de « choisir » son adversaire. Pour des raisons de sécurité et de qualité de la ressource obtenue, nous ne mémorisons que les réponses communes aux deux joueurs. Les relations ainsi acquises sont pondérées en fonction du nombre de paires de joueurs qui les ont proposées. La figure 2 présente une partie typique de JeuxDeMots où il est demandé d'énumérer ce que peut évoquer le terme *piquant*. Le joueur a fait quatre propositions.

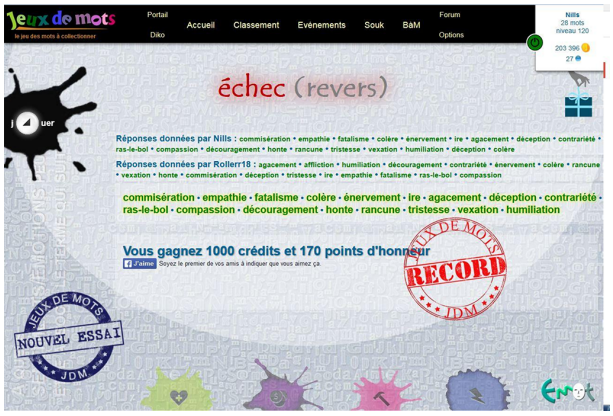


FIGURE 3 – Résultat de partie de JeuxDeMots sur le raffinement *échec>revers* avec la relation *sentiment/émotion*.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

Le résultat de la partie (figure 3) affiche la liste des termes en commun avec l'autre joueur. Les deux joueurs récoltent les gains annoncés, et, le score étant élevé, un certain nombre d'actions sont disponibles (intenter un procès, investir dans la partie ou la retenter, ajouter un ami, faire un cadeau, etc.) et constituent une récompense supplémentaire pour le joueur.

Une partie décevante peut faire l'objet d'un procès à l'initiative de celui qui vient de la jouer. Un procès consiste à rendre publiques les réponses des deux joueurs et exposer les griefs du plaignant. Les autres participants ont alors l'occasion de s'exprimer et de voter pour ou contre la partie mise en accusation. Au bout d'une semaine, le procès est clos et le gagnant (l'accusateur ou l'inculpé) reçoit un dédommagement symbolique. Les procès ont un effet pédagogique sur les joueurs, en particulier pour les relations difficiles.

Les joueurs ont également la possibilité d'offrir des cadeaux aux autres participants. Un cadeau consiste à offrir une partie sur un couple terme/relation choisi par le donateur. La motivation est ici essentiellement d'ordre social, à savoir développer une réputation positive en offrant des parties intéressantes. Le développement du réseau lexical s'en trouve accéléré et bonifié, l'immense majorité des joueurs s'abstenant d'offrir des couples terme/relation sans intérêt. L'analyse comportementale des joueurs offrant des cadeaux montre que la pression sociale (via le forum ou les procès) empêche les comportements d'« anti jeu » (par exemple offrir une partie inintéressante ou injouable).

JeuxDeMots est clairement un jeu de coopération, car en jouant bien on augmente son propre score global, mais aussi celui des autres joueurs. La compétition est induite par la présence de différents classements, sur plusieurs critères (points gagnés, nombre de parties jouées, efficacité, etc.), et renforce l'activité, et l'application de chacun à jouer le mieux possible.

Certains (contre-) jeux d'association sont dits « fermés » dans la mesure où une ou plusieurs propositions sont affichées et doivent être attribuées ou validées. En général, les parties sont plus courtes que dans JDM et jouables *via* des smartphones ou des tablettes. Ce dernier point est loin d'être anecdotique, dans la mesure où bon nombre de joueurs ont indiqué qu'ils jouaient à ces jeux dans des situations d'attente (transports en commun, etc.).

4.2 Totaki, un jeu de devinettes

Lancé début 2010, Totaki peut être considéré comme un jeu de devinettes. L'objectif était de concevoir une modalité permettant d'évaluer qualitativement le réseau lexical JDM. Cependant, le succès de ce jeu et le fait que beaucoup de joueurs cherchent à « piéger » Totaki ont également conduit à un enrichissement du réseau.

Le principe du jeu consiste à faire deviner un mot au système en lui proposant des indices. Après chaque indice, le système propose une réponse. Les indices peuvent être typés, sinon, par défaut, on les considère comme relevant simplement de l'association d'idées avec le mot à trouver. Le type d'un indice correspond à un des types de relations existant dans le réseau (par exemple « X est un véhicule »).

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850



FIGURE 4 – Partie de Totaki. Les indices (en bas à gauche), *condiment* et *piments*, ont abouti à la proposition *harissa*. Le joueur peut indiquer que c'est la bonne réponse, ou proposer un nouvel indice.

L'évaluation du réseau *via* Totaki a été présentée en 2011 [22]ainsi que l'analyse des processus impliqués [23]. Le résultat le plus notable est que Totaki devine le terme ciblé par le joueur dans environ 82 % des cas. Sous les mêmes conditions, les joueurs humains y parviennent dans 48 % des cas. Outre une évaluation des données obtenues *via* les autres jeux, Totaki permet aussi un renforcement du réseau et des détections d'incohérence par un mécanisme d'apprentissage. Depuis son lancement, plus de 276 000 parties de Totaki ont été jouées.

4.3 Jeux à votes

Le principe d'un jeu à votes est de proposer à l'utilisateur de se prononcer sur un énoncé relatif aux données du réseau. Le vote se matérialise en général par le choix d'une proposition parmi plusieurs (on parlera de « jeu à clics » si l'on s'intéresse à la modalité). Ce type de jeu a l'avantage d'être très rapide pour le joueur et de fournir une granularité très fine en termes de modification de la ressource lexicale. Suite au choix effectué par l'utilisateur, lui est notifié comment se situe son vote par rapport à ceux des autres (a-t-il voté comme la majorité ?) ; ce « retour » constitue la « récompense » qu'attend le joueur. L'objectif implicite du jeu, implicite car non annoncé explicitement aux joueurs, est qu'ils répondent honnêtement aux questions afin de faire émerger un consensus (si celui-ci est possible). Les jeux à votes sont en général de bons « contre-jeux » aux jeux ouverts, car, d'une part ils ne sont pas sujets au forçage [24] de réponses ; d'autre part, les joueurs sont relativement peu indulgents avec les réponses « discutables ».

AskIt est un jeu à clics portant sur les liens entre certains usages de termes (dits « raffinements » lorsque le terme est polysémique) et d'autres termes : par une série de questions dont les réponses possibles sont simplement « oui », « non » et « je ne

22. Lafourcade, Mathieu ; Joubert, Alain ; Schwab, Didier et Zock, Michael (2011) « Évaluation et consolidation d'un réseau lexical grâce à un assistant ludique pour le "mot sur le bout de la langue" », *Proceedings of TALN' 11*, Montpellier, France, p. 295-306.

23. Joubert, Alain et Lafourcade, Mathieu (2012) « A new dynamic approach for lexical networks evaluation », *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, Istanbul, Turkey. Disponible sur `http://www.lrec-conf.org/proceedings/lrec2012/index.html`, consulté le 6 août 2018.

24. Il s'agit de la tendance qu'a souvent le joueur à vouloir répondre coûte que coûte, qui le conduit à donner des réponses qui s'écartent de la consigne.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

sais pas », il permet un enrichissement et une consolidation du réseau, en particulier en ce qui concerne les raffinements des termes polysémiques.



FIGURES 5 et 6. Exemples de deux écrans consécutifs de AskIt.

AskIt est un jeu de consensus fermé et rapide. L'analyse des motivations des joueurs montre que l'excitation/frustration suite à une partie est liée à l'espérance/déception de se situer ou non dans la majorité des réponses. Les points gagnés font progresser une barre d'état, et l'objectif de la remplir pour passer au niveau supérieur est une incitation à répondre correctement. Le taux de « trollage » (joueurs répondant volontairement quelque chose d'erroné) a été estimé à moins de 1 % sur l'ensemble de plus de 300 000 parties jouées à ce jour.

AskIt est un jeu relativement addictif :

— Le système « relance » le joueur qui arrêtera le jeu sur une question simple qu'il a eu le temps de lire (environ 2 à 3 secondes) et dont il connaît en général la réponse. Sentiment de satisfaction (on a fini une partie)/frustration (on arrête sur une partie en cours à laquelle on peut répondre d'un simple clic) ;
— Il peut être amusant, voire accrocheur, en particulier sur les termes polysémiques (ex. : « est-ce que le chien [tarot] peut aboyer ? ») ;
— Les parties sont très courtes : elles ne durent que quelques secondes. Cependant, la séquence de jeu peut être longue (plus d'une demi-heure) de l'aveu même des joueurs ;
— Les réponses données sont spontanées : est-ce mieux si les joueurs ne sont pas linguistes ? On obtient une connaissance générale commune et non une connaissance de spécialistes. Cependant, si la question est trop complexe ou ambiguë, elle peut entraîner un grand nombre de non-réponses [25] ;
— Le résultat (« qu'en ont pensé les autres joueurs ? ») est donné immédiatement : en comparant sa réponse avec celles des autres, le joueur est valorisé si sa réponse est conforme à la majorité (« j'ai raison »), mais il l'est également si elle s'en distingue : il peut alors avoir le sentiment d'avoir fourni une réponse astucieuse ou réfléchie (« je suis plus malin qu'eux »).

Le jeu LikeIt concerne les relations de type polarité *positive/négative/neutre*, simplifiées par la question fermée « Est-ce que vous aimez l'idée de » suivi d'un terme. Les

25. Le joueur « passe », clique sur « je ne sais pas », ou encore recharge la page pour avoir une nouvelle partie.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

seules réponses possibles sont « oui », « non » et « neutre ». Ce jeu permet ainsi d'enrichir les termes du réseau en les caractérisant par une polarité. L'exploitation préliminaire de ces seules données, dans le cadre d'une désambiguïsation lexicale, semble montrer que la polarité prise isolément permet de sélectionner le bon usage d'un terme en contexte dans 50 % des cas. Ce type de données peut aussi être exploité en analyse d'opinion.

LikeIt est un jeu de consensus, qui possède un certain nombre des qualités définies précédemment : comme AskIt, LikeIt est un jeu addictif, à parties courtes et relance immédiate :

— Les réponses sont très simples : « Oui, j'aime l'idée », « Non, je n'aime pas l'idée », et « ça ne me fait ni chaud ni froid » : le joueur n'a pas l'impression de renseigner la base, mais la variété du vocabulaire est telle qu'il a plutôt tendance à voir ce jeu comme un sondage ;

— Certains termes donnent lieu à des réponses très contrastées (ex. : le terme *bloc opératoire* est neutre, voire positif dans l'absolu, mais négatif lorsqu'on est personnellement concerné) ; l'avis d'un même joueur peut évoluer dans le temps, en fonction du contexte (ex. : le terme *baccalauréat* inspire un sentiment négatif quand on est lycéen, mais est perçu bien plus positivement lorsqu'on l'a obtenu).

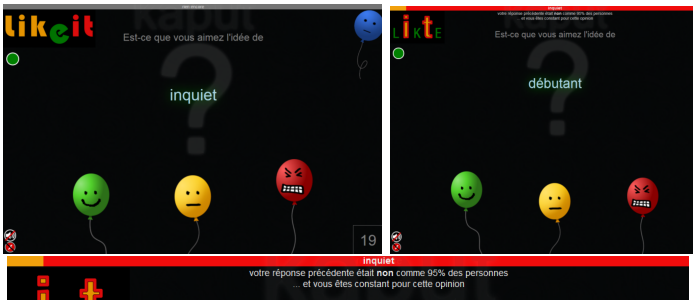


FIGURE 7, 7bis et 8. Exemple de deux écrans consécutifs de LikeIt. La réponse donnée dans l'écran de gauche se trouve en haut de l'écran de droite (*cf.* zoom en dessous), fournissant ainsi un retour direct au joueur avec une relance immédiate.

La sélection d'un terme à proposer au joueur est effectuée pseudo-aléatoirement dans le réseau lexical de JeuxDeMots. Un terme déjà polarisé partiellement positivement ou négativement (mais pas uniquement neutre) est sélectionné au hasard ainsi qu'un autre, parmi tous les termes qui lui sont reliés. Ainsi, cet algorithme très simple (de marche aléatoire) effectue une propagation au sein du graphe entre termes potentiellement intéressants pour l'acquisition de valeurs de polarité. Une sélection strictement aléatoire au sein du réseau lexical donnerait une proportion trop importante de termes neutres, ce qui diminuerait l'intérêt du jeu. En trois mois, plus de 24 000 termes ont été polarisés, ce qui représente un total dépassant les 150 000 clics. En août 2020, plus de 900 000 termes ont été polarisés.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

5 Analyse de la ressource obtenue

Il est possible, *via* l'analyse des données du réseau JeuxDeMots, de déduire les difficultés que peuvent éprouver les joueurs face à certains types de relations lexico-sémantiques.

5.1 Quelques chiffres, et quelques consignes de jeu problématiques

Depuis le lancement du projet JeuxDeMots, le réseau lexical a acquis plus de 300 millions de relations et 4,1 millions de termes (en toute généralité des objets lexicaux). Il y a environ 100 types de types de relations lexico-sémantiques dans JDM, la plupart d'entre elles pouvant être jouées ; celles qui présentent un intérêt ludique trop limité pour être jouées (comme les lemmes ou les *POS*) peuvent faire l'objet d'une contribution directe *via* Diko, l'outil contributif du projet JeuxDeMots [26].

Tableau 1. Nombre d'occurrences des principales relations du réseau lexical obtenues au 30 mars 2018.

104 831 733 associations libres	86 118 r_domain (domaines)	1 025 013 synonymes
4 540 364 hyperonymes	83 095 antonymes	823 225 hyponymes
4 366 151 parties (méronymes)	3 795 953 tout (holonymes)	1 563 686 locutions
1 403 447 agents (verbes)	73 084 patients (verbes)	553 703 lieux typiques
26 616 instruments (verbes)	464 911 caractéristiques typiques	

On notera que 65 041 lemmes et 88 758 catégories morphosyntaxiques ont été obtenus de façon automatique ou contributive, donc hors-jeux.

Comme nous l'évoquons plus haut, ces informations ne sont pas amusantes à jouer et présentent trop peu de variations dans une entrée pour avoir un intérêt ludique. Il est donc nécessaire de les entrer manuellement ou par tout autre moyen (semi) automatique.

Les joueurs ne fournissent pas directement les raffinements sémantiques, mais ont le loisir de jouer à « gloser » les termes, c'est-à-dire à proposer des mots ou expressions représentatifs des usages possibles. Les joueurs trouvent cette tâche particulièrement difficile, mais intéressante. Parfois, l'introduction de périphrases semble être la seule réponse possible, si aucun hyperonyme ou synonyme ne semble convenir.

Enfin, certaines relations posent plus de problèmes aux joueurs.

Certains ont par exemple du mal à identifier clairement le type d'une relation, pour les types de relations qui admettent une relation inverse ou symétrique. L'analyse de certains fils de discussion liés au projet JDM montre que le problème n'est pas d'ordre linguistique, mais plutôt lié à la connaissance du monde. Par exemple, entre fièvre et malade, identifier ce qui relève de la cause ou de la conséquence semble difficile. De même, il est parfois difficile de choisir entre producteur et cause (c'est le problème de l'œuf et de la poule) d'une part, et produit et conséquence d'autre part (le lait est-il un produit ou une conséquence de la vache ?).

26. `http://www.jeuxdemots.org/diko.php`.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

La distinction entre instance et hyponyme est une autre difficulté : pour beaucoup de joueurs, les hyponymes semblent inclure les instances (mais pas l'inverse). Par exemple, *Jolly Jumper* est vu aussi bien comme une instance de cheval que comme un hyponyme. Par contre, *pur-sang* n'est (à juste titre) considéré que comme un hyponyme.

La distinction entre le rôle agentif et les étapes de construction est également problématique : le rôle agentif est le ou les verbes de construction de l'entité/artéfact (par exemple, *construire* pour *maison*, ou *écrire/éditer* pour *livre*[27]). L'implication agentive est le ou les verbes présidant aux étapes de construction (par exemple pour un *gâteau* : *mélanger, enfourner, cuire, démouler*, etc.). La distinction entre la sémantique de chacun des deux types de relations n'est pas immédiatement intuitive pour les joueurs, et il s'ensuit un nombre important de réponses erronées.

En ce qui concerne la distinction entre magn (« plus intense ») et bon (mélioratif), on s'aperçoit que l'axe d'intensité pour un terme est souvent parfaitement subjectif, et les « plus intenses » (dont un exemple canonique est « *fièvre* a comme plus intense *forte fièvre, fièvre de cheval* », etc.) incluent les mélioratifs (« *excellent repas* donné comme plus intense que *repas* »). Un autre exemple est de considérer *belle maison* comme plus intense que *maison* (ce qui n'est pas strictement faux) mais qui correspond plutôt à un mélioratif.

Enfin, pour le type de relation « Termes dérivés morphologiquement » sont attendus des termes appartenant à la même famille lexicale de façon dérivationnelle. Or, il arrive que les résultats soient étonnants ; par exemple, pour le substantif *nuit*, nous obtenons : *nuitée, nocturne, noctambule, nuitamment, minuit, noctambulisme, chemise de nuit, noctuelle, nuitamment, noctule, bonnet de nuit, nuire, nuisible* ». Manifestement *nuire* et *nuisible* sont des intrus, et *nocturne* est un mot étymologiquement apparenté, pour lequel le type de relation adéquat existe d'ailleurs. Une évaluation par échantillonnage semble indiquer qu'environ 5 % des termes associés par cette relation sont des erreurs (lesquelles sont parfois largement partagées parmi les locuteurs). Il s'agit d'une relation clairement lexicale (et non sémantique).

5.2 Quelques particularités du réseau JeuxDeMots

Un raffinement[28] est un terme comme les autres et peut être lui-même raffiné. Par exemple, pour le terme *frégate*, nous avons deux niveaux de raffinement, *frégate>navire* étant lui-même raffiné :

27. Pustejovsky, James (1991) « The Generative Lexicon », *Computational Linguistics*, MIT Press, Cambridge, USA, 17 (4), p. 409-441.

28. Chapuis, Maxime et Lafourcade, Mathieu (2017) « Identifying Polysemous Words and Inferring Sense Glosses in a Semantic Network », *Proceedings IWCS : International Conference on Computational Semantics*, 12th International Conference on Computational Semantics, September 19-22, Montpellier, France. Disponible sur `https://hal.archives-ouvertes.fr/lirmm-01763423`, consulté le 3 août 2018.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

frégate	frégate > navire frégate > oiseau frégate > automobile frégate > bicorne frégate > giton	frégate > navire > moderne frégate > navire > ancien

Le verbe *louer* constitue un autre exemple de raffinement hiérarchique :

louer	louer > louage louer > admirer	louer > louage > donner à louage louer > louage > prendre à louage

La hiérarchie de raffinement peut être utilisée comme un arbre de décision quand le contexte d'une phrase n'est pas suffisant pour désambiguïser complètement le terme, comme dans les exemples suivants :

Lors de la bataille navale deux frégates ont coulé. (frégate > navire)

Il loue un appartement dans le centre ville. (louer > louage)

Par contre les petits textes suivants fournissent assez de contexte pour choisir les raffinements les plus précis :

La frégate a été détruite par un missile. (frégate > navire > moderne)

Il a des problèmes avec le propriétaire de l'appartement qu'il loue. (louer > louage > prendre à louage)

Les pondérations sur les relations du réseau peuvent être négatives, ce qui permet plusieurs choses :

— Représenter des exceptions, par exemple : *voler agent < 0 autruche* ;
— Représenter des informations négatives, par exemple : *accoucher r_agent < 0 homme > être humain de sexe masculin* ;
— Enfin, mémoriser des erreurs présentant un intérêt, qu'elles soient faites par les joueurs ou des processus automatiques : par exemple, beaucoup de gens pensent qu'une *araignée* est un *insecte*, ce qui est faux, ou confondent *Dalida* (chanteuse) et *Dalila* (figure biblique).

Les relations négatives peuvent être exploitées lors d'une analyse automatique de texte afin d'éliminer des interprétations impossibles. Par exemple, pour la phrase « *La frégate s'est envolée au premier bruit.* », le raffinement *frégate>oiseau* peut *s'envoler* mais comme c'est impossible (relation négative) pour le raffinement correspondant au navire, c'est le raffinement *frégate>oiseau* qui est sélectionné.

Enfin, les relations peuvent être annotées avec des termes du lexique. Par exemple, la relation *rougeole r_cible adulte* peut être annotée avec les termes *rare* et *dangereux*. Ou encore, l'annotation *peu pertinent* concerne la relation *souris has_part atome* et la signale comme non-informative. Outre le fait de pouvoir ajouter des

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

méta-informations aux relations, les annotations peuvent aussi être des éléments déterminants pour l'application de mécanismes d'inférence.

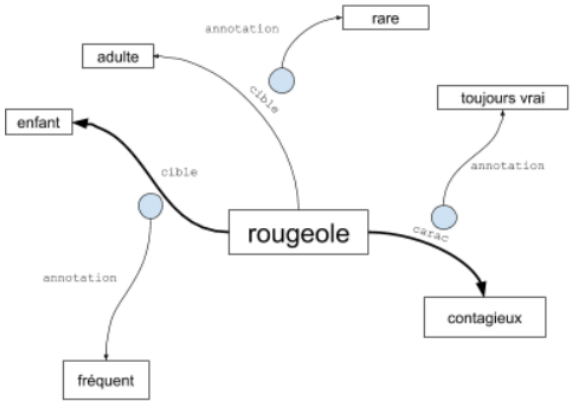


FIGURE 9. Exemple d'annotations dans le réseau JeuxDeMots.

5.3 Erreurs et inférences

Un des biais majeurs de JDM veut que la consigne soit potentiellement forcée – « il faut absolument trouver et mettre des réponses ». Le joueur a bien entendu la possibilité d'indiquer qu'il n'en trouve aucune, mais il arrive qu'il préfère indiquer ce qui lui semble le plus approchant. C'est ainsi que pour des spécifiques de semaine ont été proposés les jours de la semaine : *lundi, mardi,* etc., qui sont plutôt des spécifiques (ou hyponymes) de *jour* ; comme hyponyme de semaine, on aurait attendu *semaine de vacances*, *semaine sainte*, ou encore *semaine prochaine*. Les relations erronées issues de ce type de phénomène de forçage ont été évaluées (par échantillonnage) à environ 0,1 % des relations. Il est à noter que la présence des contre-jeux vient contrebalancer largement ce biais propre aux jeux ouverts. Toutefois, nous avons remarqué que les espaces d'échanges entre joueurs (les procès et le forum) sont des lieux privilégiés pour éclaircir la sémantique des relations quand elle n'est pas claire pour le joueur, qui ainsi s'améliore progressivement au fil de sa pratique et de l'interaction avec ses pairs.

De façon plus précise, nous avons évalué les performances globales des joueurs. Dans JDM, chaque partie en stock est signée (nous savons quel joueur en est le créateur) ; de plus, il est possible de calculer de façon incrémentale le nombre moyen de termes en commun entre les joueurs deux à deux (on parlera d'« efficacité »). En mars 2018, l'efficacité moyenne des joueurs est de 2,85 (c'est-à-dire qu'ils ont en moyenne 2,85 termes en commun avec l'autre joueur lors d'une partie). Les meilleurs joueurs ont une efficacité qui se situe entre 10 et 11.

Une évaluation statistique globale a permis d'estimer à 0,1 % le nombre de relations manifestement erronées, et à moins de 0,4 % celles qui sont discutables (données évaluées en juillet 2018). Environ 86 % des relations erronées sont repérables automatiquement. Il s'agit par exemple des doubles hyperonymes (A a comme hyperonyme

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

B et B a comme hyperonyme A) qui, *modulo* la polysémie, posent problème (exemple faux : « un chat est un matou/un matou est un chat », car la relation entre les termes relève plutôt de la synonymie ; exemple juste, à cause de la polysémie : « une plante a un pied, un pied a une plante »). Environ 14 % des relations erronées ne sont pas repérées automatiquement, soit parce qu'il s'agit d'exceptions, soit parce que les informations disponibles pour détecter l'erreur sont en nombre insuffisant. Il est à noter que les faux négatifs sont estimés à environ 12 % (c'est-à-dire que les approches automatiques pensent qu'il s'agit de relations fausses alors qu'elles sont plutôt justes). Nous n'invalidons donc pas automatiquement les relations détectées comme fausses, car cela nous amènerait à perdre beaucoup de relations valides. Environ 3 % des relations fausses concernent des relations de faible poids, et ces erreurs tendent à se résorber naturellement avec les contre-jeux.

Nous avons également constaté qu'une relation sémantique nouvelle introduite par un couple de joueurs dans JDM permet de générer environ 57 autres relations sémantiques grâce aux mécanismes automatiques d'inférences (de type déductives pour l'essentiel).

6 Conclusion

Nous avons montré l'intérêt du projet JeuxDeMots, fondé sur des jeux à finalité (GWAP), mais également des processus automatiques, qui ont conjointement permis d'acquérir facilement et à moindre coût des ressources lexicales et sémantiques de qualité pour le français. À la conception traditionnelle d'un jeu sur Internet s'ajoute la difficulté d'avoir une boucle ludique covariante avec la boucle d'acquisition de données. Cette méthode de construction est relativement récente puisqu'elle n'est réalisable que depuis l'émergence d'Internet.

Nous avons observé que la collecte d'informations relevant aussi bien du lexique que du sens commun se faisait aisément *via* des jeux, et ce avec une qualité très satisfaisante. Plus l'information demande une expertise linguistique poussée, plus les résultats sont aléatoires. Concernant la constitution d'un réseau lexical, peu de données exigent une compétence linguistique qui dépasserait l'intuition du locuteur moyen. Il n'en serait certainement pas de même s'il s'agissait, par exemple, de constituer une banque d'arbres d'analyse syntaxique, ou encore des alignements de traduction. Nous avons détaillé un outil permettant la collecte de polarité, qui est d'un abord très simple : l'activité est à granularité très fine, et l'intuition des locuteurs correspond aux attendus. Il s'agit là d'un exemple typique où une activité ludique remplace avantageusement un travail de collecte manuelle, aussi bien quantitativement que qualitativement.

L'évaluation de la ressource peut se faire également de façon efficace *via* une activité ludique. De fait, ce sont les joueurs eux-mêmes qui s'en chargent, *via* les contre-jeux au jeu principal : en répondant à des questions fermées, les joueurs valident ou invalident les données lexico-sémantiques recueillies *via* le jeu principal. Par ailleurs, un des avantages majeurs de ce genre d'approche est que l'échantillonnage est effectué implicitement par les joueurs eux-mêmes : *via* les cadeaux qu'ils se font, les joueurs renseignent prioritairement ce qui leur semble intéressant, et la ressource se construit donc principalement sur des termes présentant de l'intérêt. Et *via* le jeu Totaki, qui

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850

réalise aussi une évaluation, ce sont encore les joueurs qui échantillonnent les termes à évaluer.

Published in : Cécile Meynard et Thomas Lebarbé (dir.), *Le Crowdsourcing : Partager, enrichir et publier des sources patrimoniales*, Editions des archives contemporaines, ISBN : 9782813003850, doi : 10.17184/eac.9782813003850